DOCKET NO: 0010-0929-0X

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

IN RE APPLICATION OF	:
YUTARO KANEKO, ET AL.	: EXAMINER: WILSON, MICHAEL
SERIAL NO: 09/087,513	:
FILED: MAY 29, 1998	: GROUP ART UNIT: 1632
FOR: METHOD FOR INDUCING IMMUNITY TO VIRUSES	:

DECLARATION UNDER 37 C.F.R. § 1.132

COMMISSIONER FOR PATENTS
ALEXANDRIA, VIRGINIA 22313

SIR:

Now comes Danuta Kozbor, who deposes and states that:

(1) I am a graduate of Queen's University, Canada and received a Philosophy Degree in Microbiology and Immunology degree in the year 1982 as shown in my C.V. attached hereto as Exhibit 1.

(2) I have been employed by Roswell Park Cancer Inst, for 2.5 years as a researcher in the field of Immunology.

(3) I understand the English language or, at least, that the contents of this Declaration were made clear to me prior to executing the same.

(4) I understand that the U.S. filing date of the above-identified application, U.S. application serial No. 09/087,513, is May 29, 1998.

(5) I have read and am familiar with the contents of the above-identified application.

(6) The field of the invention described in the above-identified application is methods and compositions for inducing immunity against the human immunodeficiency virus (HIV).

(7) I have read and am familiar with the subject matter of Claims 14, 15, 19, and 21-36 of the above-identified application. The text of the claims of the above-identified application is attached hereto as Exhibit 2.

(8) I have read and am familiar with the contents of the Official Action dated June 25, 2003 in the above-identified application.

(9) I have read and am familiar with the contents of the Official Action dated February 13, 2004 in the above-identified application.

1. Naming History of HIV

(10) The Human Immunodeficiency Virus-1 (HIV 1) was discovered as the virus causing Acquired Immunodeficiency Syndrome (AIDS) in 1983-1984 by two different groups. The first group was led by Luc Montagnier at the Pasteur Institute in France. The second group was led by Robert C. Gallo at the National Cancer Institute in the U.S. The French group named the virus as lymphoadenopathy virus (LAV), while the U.S. group named it human T cell lymphotropic virus type III (HTLV-III).

(11) Later, these two viruses were found to belong to the same family and re-named as human immunodeficiency virus type 1 (HIV-1) as a family name. The virus isolated by the French group was renamed as the HIV 1 lav strain and the virus isolated by the U.S. group was renamed as the HIV 1 IIIb strain.

(12) The HIV-1 virus can mutate its sequence spontaneously and very rapidly. Such a sequence change is usually caused by a point mutation in the variable regions. In fact, many strains are obtained from various patients, and each isolated virus possesses slightly

different amino acid sequence as compared to the others. In addition, because of the nature of the HIV-1 virus, a strain can mutate spontaneously during culturing in a laboratory over a period of years. Therefore, there are many kinds of slightly different sequences reported in published scientific articles and in the databases.

(13) Thus, the term "HIV-1 IIIb" refers to a prototypic virus isolated in the early years of HIV research. In fact, the term "HIV-1 IIIb variant" was used to refer to family strain derived from the laboratory strain originally isolated by the group led by Dr. Gallo.

(14) Reitz et al. (*AIDS Res. Hum. Retroviruses* 10 (9), 1143-1155 (1994)) clearly state that HIV-1 strains BH10, BH8, HXB2 and HXB3 are isolated clones from HIV-1 (HTL-VIIIB) (see abstract and page 1145, right column, lines 7-9). A copy of Reitz et al. is attached hereto as Exhibit 3. Furthermore, a sequence comparison reveals these typical HIV-1 IIIb variants have quite similar sequences because they originally derived from Dr. Gallo's isolate. Exhibit 4 attached hereto presents the GP160 sequences from various HIV-IIIb variants.

(15) There are small sequence differences between these strains, but these differences are not significant for practicing the invention described in the above-identified application. In the invention described in the above-identified application, the essence is inducing a CTL response against conserved region by deleting the V3 loop, immunodominant region. Therefore, there is no need to practice the examples specifying all varieties of HIV-1 strains' sequences nor one specified sequence which can easily vary. Rather, it is sufficient to say, for example, "Δ V3 loop" or "V3 loop deletion." Typical sequences for HIV-1 IIIb such as HXB2 or BH10 had been available and well-recognized in the field of the invention at the time of filing date of the above-identified application. For example, the following table lists the available date for a part of the HIV-1 IIIb strain sequences data in the database.

Isolates of HIV-1 IIIb	Genebank Accession no.	Available in NCBI Since:	Source created in:
HTLV-IIIB	AAA76690	Feb 15 1995	1987 isolate
BH10	P03375	Apr 23 1993	Jul 21, 1986
HXB2	P04578	Apr 23, 1993	Aug 13, 1987
HXB3	P04624	Apr 23 1993	Aug 13, 1987

(16) The notation "Δ" means the deletion or destruction of the indicated component. Thus, that notation was also well-known, common technical language among scientists in the field of the invention at the time the present application was filed. Thus, it would have been clear to those skilled in the field of the invention at the time the present application was filed that the expression "Δ V3 loop" means that the V3 loop was deleted.

2. History and Relationship Between the Variable Region V3 and the V3 Loop

(17) In 1987, Modrow et al. (*Journal of Virology* 61 (2), 570-578 (1987)) reported a sequence comparison of the envelope protein, which included the BH10 and LAV isolates of HIV-1. A copy of Modrow et al. is attached hereto as Exhibit 5. At that time, the term "HTLV" was still used for HIV-1. Therefore, some of HIV strains were referred to as HTL-VIII as termed by Dr. Gallo. In Modrow et al., the comparison revealed conserved regions, variable regions and cystein residues (see FIG 1). Note that cystein residues do not vary among several strains.

(18) Modrow et al. named conserved regions C1 to C6, and variable regions V1 to V5 (see FIG 2). In the figure shown in that reference, variable region V3 is from amino acid positions 300 to 321 in the BH10 isolate according to Table 2 or is from amino acid positions 300 to 329 in the BH10 isolate according to Fig. 1. After the publication of Modrow et al., the term "variable region 3 (V3)" became common language in the area of HIV research.

The number, 300 or 321/329, may slightly vary strain by strain (report by report) because of virus mutability or a deficiency in the numbering standard for such variation. See Korber et al. 1998 (*HIV Molecular Immunology Compendium* 1998, Review articles, "Numbering Positions in HIV Relative to HXB2CG"), a copy of which is attached hereto as Exhibit 6. This difference was also recognized by those skilled in the field of the invention at the time the above-identified application was filed.

(19) In 1990, Leonard et al. (*The Journal of Biological Chemistry* 265 (18), 10373-10382 (1990)) reported conformational evidence for envelope protein: several loop formations closed by two cystein residues. A copy of Leonard et al. is attached hereto as Exhibit 7. Leonard et al. revealed loop formation bridged by SH-group of cystein residues (disulfide bond) and found variable regions V1, V2, V3 and V4 are included in each loop. Leonard et al. showed conformational model indicating variable regions Vl to V5 in Figure 7, and variable region V3 is indicated on the loop formed by two cystein residues. This loop formation, protruding from the envelope protein surface, was quite understandable to those skilled in the field of the invention at the time the above-identified application was filed, since these variable regions are dominant epitopes.

(20) In Figure 6, variable region V3 is indicated from amino acid positions 270 to 298, those omitting 30 amino acids of signal peptide, therefore, adjusting to BH-10 isolate sequence, this is equivalent to amino acid positions 300 to 328 in BH10 isolate (see sequential analysis sheet). Variable region V3 is located within the loop formed by two cystein residues at the amino acid positions 296 and 301 in the BH10 isolate.

(21) This is summarized by Emini et al. 1992 (*Biotechnology Series*, 20, 309-326 (1992)). A copy of Emini et al. is attached hereto as Exhibit 8. In Emini et al., Figure 13-1 shows a loop formation corresponding to variable region V3. As noted above, cystein residues forming the V3 loop do not vary among the strains having different sequences;

therefore, the skilled person in the field of the invention can easily recognize the V3 loop, which is formed by two cystein residues located at 296 and 331 in HIV-1 IIIb variants. Even when this amino acid position numbering differs because of strain differences, one skilled in the field of the invention at the time the above-identified application was filed could have found the V3 loop by finding the corresponding two cystein residues or by finding the variable region V3 using a homology comparison.

(22) This is why the indicated position for "V3" which means "variable region V3" sometimes does not match the exact position for "V3 loop," such as described in Back et al. (*Journal of Virology*, Nov. 1993, pp. 6897-6902), a copy of which is attached hereto as Exhibit 9. This loop formation, i.e., the V3 loop, became common terminology in the field of the invention. After that, researchers commonly refer to the V3 loop as almost the same meaning as for variable region V3. This structural knowledge was well-accepted in the field of the invention prior the filing date of the above-identified application. See, Emini et al.

(23) Wyatts published several articles describing loop-deletion mutants to investigate functional analysis of the envelope protein, and he always used the expression "V3 deletion" to mean that the natural sequence of the third variable loop was replaced with the sequence GAG. His publications support the observation that those skilled in the field of the invention recognized that a V3 deletion mutant constructed by Wyatts (Dr. Sodoroski's research group) included a GAG in place of the deleted portion of the sequence. This terminology was widely known to those skilled in the field of the invention at the time the present application was filed.

(24) Even if the specification of the above-identified application does not refer to Wyatts but refers to a V3 deletion mutant, for example, those skilled in the field of the invention would have understood at the time the present application was filed what kind of construct was to be made, based on the knowledge described above. Even when just an

indication to delete the V3 loop is given, the skilled person can understand which portion is to be deleted to obtain desired effect of the invention, and can find a suitable position to be deleted (note: 297-329 is apparently included in the V3 loop of IIIb variants such as BH10 or HXBc2). Also, to maintain the correct conformation of the envelope protein, the skilled person would understand closing the edge after deletion of V3 loop, and the easiest connecting peptides might be those consisting of small amino acid such as glycine or alanine.

(25) The V3 loop is located between two cystein residues located at amino acid positions 296 and 331 in HIV-1 IIIb variants. These cystein residues are not located around amino acid positions 99-110, which are the amino acid position calculated by, incorrectly, considering 297-329 as referring to the nucleotide position in the encoding nucleic acid, as suggested by the Examiner.

(26) Also, the biotechnology tools for genetic engineering were sufficient for constructing a V3 loop deletion mutant which was not explicitly described by Wyatts or Sodoroski. Those skilled in the field of the invention can find suitable primers based on the sequence data as a template. To delete V3 loop formation, one only has to design a suitable restriction enzyme to cut out V3 region and then insert suitable bridge between the cut positions.

(27) Based on the foregoing, the amino acid sequence of the envelope glycoprotein of HIV was well-known to those skilled in the field of the invention at the time the present application was filed in the U.S. In particular, the location of the variable region of the amino acid sequence was well-known to those skilled in the field of the invention. Thus, those skilled in the field of the invention at the time the present application was filed in the U.S would have appreciated that the description of the numbers 297 to 329 in the specification of the above-identified application was referring to the amino acid sequence of the protein and not to the nucleic acid sequence encoding that protein.

(28) The undersigned petitioner declares further that all statements made herein of his own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of this application or any patent issuing therefrom.

(29) Further deponent saith not.

Danuta Kozbor

July 11, 2004
Date

I:\ATTY\JK\101003US-132DEC.DOC

CURRICULUM VITAE

Name Danuta Kozbor, Ph.D.

Address
Department of Immunology
Roswell Park Cancer Institute
Elm & Carlton Streets
Buffalo, NY 14263

Phone: (716) 845-8668
E-mail: danuta.kozbor@roswellpark.org

Citizenship USA

Education

1967 - 1971	B.Sc. (Natural Sciences), VII Lyceum, Krakow, Poland.
1971 - 1976	M.Sc. Jagiellonian University, Inst. of Molecular Biology, Krakow, Poland.
1977 - 1978	First year of Ph.D., Karolinska Institute, Dept. of Tumor Biology, Stockholm, Sweden.
1979 - 1982	Ph.D., Queen's University, Dept. of Microbiology and Immunology, Kingston, Ontario, Canada.

Awards and Honors

1980-1982	National Cancer Institute of Canada Studentship, Queen's University, Kingston, Ontario (Supervisor: Dr. John C. Roder).
1982-1986	National Cancer Institute of Canada Fellowship, The Wistar Institute, Philadelphia, PA (Supervisor: Carlo M. Croce).
1988-1990	The W. W. Smith Charitable Trust Grant: "Effect of HIV on CD3/TCR Function and CD4 Antigen Expression".

1990-1996	National Institute of Child Health and Human Development R01HD27107-06: "Cellular Activation in Primary Perinatal HIV Infection".
1996-2001	National Institute of Allergy & Infectious Diseases R01AI/HD39148-07: "T-cell Helper and Cytotoxic Activity in Pediatric AIDS."
1994-2000	Japan Health Science Foundation: "Study of treatment and prophylaxis of HIV infection by induction of cytotoxic T lymphocytes (CTLs)".
April 10-15, 2002	Speaker at Keystone Symposium on ***Gene-Based Vaccines: Mechanisms, Delivery Systems and Efficacy***, talk entitled: "Adjuvanted Mucosal Delivery of Microparticle- and Liposome-Formulated DNA Vaccines".
June - Sept. 2000	Visiting Professor, Division of Viral Immunology, Center for AIDS Research, Kumamoto University, Kumamoto, Japan.

Academic Appointments

1976 - 1977	Research Assistant, Department of Physiology, Medical Academy, Krakow, Poland.
1982 - 1985	Postdoctoral Fellow of the National Cancer Institute of Canada in The Wistar Institute, Philadelphia, PA.
1985 - 1989	Associate Scientist, The Wistar Institute, Philadelphia, PA.
1989 - 1991	Assistant Professor, Fels Institute for Cancer Research and Molecular Biology, Temple University School of Medicine, Philadelphia, PA.
1991 - 1996	Assistant Professor, Department of Microbiology and Immunology, Jefferson Cancer Institute, Thomas Jefferson University (TJU), Philadelphia, PA.
1996 - 1998	Associate Professor, Center for Neurovirology and NeuroOncology Department of Neurology, Allegheny University of the Health Sciences, Philadelphia, PA.
1998 - 1999	Associate Professor, Department of Microbiology and Immunology Thomas Jefferson University, Philadelphia, PA.

1999 - 2002	Professor, Center for NeuroVirology and Cancer Biology College of Science and Technology, Temple University, Philadelphia, PA.
2002 - present	Associate Professor, Department of Immunology, Roswell Park Cancer Institute, Buffalo, NY.

Other Professional Activities

1985-1985	Invited Instructor: International Workshop Course on Hybridomas, National Institute of Immunology, New Delhi, India (Dec. 2-14) organized by World Health Organization.
1986-1987	Consultant on the Hybridoma Technology, World Health Organization (WHO Regional Office for Southeast Asia, New Delhi, India).
1986-1988	Consultant on the Hybridoma Technology, Dow Chemicals, Midland, MI.
88-1989	Consultant on the Hybridoma Technology, DuPont Company, Wilmington, DE.
1994-1999	Ph.D. Thesis Advisory Committee, University of Pennsylvania, Philadelphia, PA.
1996-1999	Clinical Trials in Zambia: "A phase II evaluation of curdlan sulfate bolus infusion alone in combination with quinine in patients infected with malaria".
1995-2002	Investigator, Philadelphia Pediatric AIDS Clinical Trials Unit (PI Stuart E. Starr, M.D.)
1999-2002	Investigator, Penn Center for AIDS Research, Philadelphia, PA.
999-present	Editorial Board Member, Clinical & Diagnostic Laboratory Immunology.
1999-present	NIH/NIAID Special Emphasis Review Panel on HIV Vaccine Research & Design
2002	Reviewer, NIH-NIAID-DAIDS-03-12, "HIV Vaccine Design and Development Teams"
2003	Reviewer, NIH-NIAID-DMID-03-04, " Food and Waterborne Diseases Integrated Research Network"

Research Projects (ongoing and completed in last 3 years)

Completed

R01AI/HD39148 (Danuta Kozbor, PI) 09-01-1996 to 08-31-2001.

National Institute of Allergy & Infectious Diseases:
"T-cell helper and cytotoxic activity in pediatric AIDS"

The major goal of this study was to analyze the relationship between the profile of cytokine production and HIV-specific CTL responses during progression to AIDS in HIV-infected children born to HIV-infected mothers in order to understand the correlates of protective immunity to HIV.

R21AIDE48370-02 (Danuta Kozbor, PI) 07-01-2000 to 06-30-2003
National Institute of Allergy & Infectious Diseases (No cost extension)
"Mucosal immunity to HIV env by oral vaccination"

The long-term objective of this research project is to develop immunization strategies for inducing high levels of HIV envelope glycoprotein-specific immunity in systemic and mucosal tissues by an oral vaccine with PLG-encapsulated plasmid DNA in a murine animal model.

Japan Health Science Foundation (Danuta Kozbor, PI) 04-01-2001 to 03-31-2004
"Immunization approaches to enhance mucosal immunity to HIV antigens"

Active

R21 AI54163-01 (Danuta Kozbor, PI) 06-01-2002 to 05-31-2005
National Institute of Allergy & Infectious Diseases
"Vaccine with altered HIV antigens"

The major goal of this study is to increase immunogenicity of the HIV Gag/Pol fusion protein by redirecting immune responses to subdominant and conserved epitopes by DNA vaccines in a murine model system.

Patents

1984 November 11.1984; Entitled "Human Hybridoma Fusion Partner for Production of Human Monoclonal Antibodies." (U.S. Patent Serial No. 673,370).

Other University Positions

1989-1991	Supervisor, Flow Cytometry Facility, Fels Institute for Cancer Research and Molecular Biology, Temple University School of Medicine
1991-1996	Supervisor, Jefferson Cancer Center (JCC) Flow Cytometry Facility
1991-1996	Member, Committee on Graduate Program in Immunology
1991-1996	Member, JCC Glassware and Tissue Culture Committee
1993-1996	Member, Institutional Biosafety Committee, TJU
1995-1996	Coordinator, Graduate Program Course on "Tumor Immunology"
1999	Member, Institutional Biosafety Committee, TJU
2002-present	Member, Institutional Biosafety Committee, Roswell Park Cancer Institute

Editorial Activities

1987-1989	Manuscript Referee, Hybridoma
1988	Manuscript Referee, Procedins of the National Academy of Sciences
1990	Manuscript Referee, Cancer Research
1993	Manuscript Referee, Journal of Acquired Immunodeficiency Syndromes
1993	Manuscript Referee, Analytical Biochemistry
1995	Manuscript Referee, International Immunology
1995	Manuscript Referee, Clinical Immunology and Immunopathology
1999-present	Editorial Board Member, Clinical & Diagnostic Laboratory Immunology

Professional Organization Memberships and Positions

1985	The New York Academy of Sciences
1994	American Association for the Advancement of Science

Teaching Experience

Teaching Positions

1990-1991	Temple University School of Medicine:
	Teacher, Medical Immunology and Molecular Biology for graduate students
1991-1999	Thomas Jefferson University:
1991-1994	Teacher, Microbiology-Immunology, Sophomore medical students
1991-1996	Teacher, Fundamentals of Immunology, graduate students
1993-1996	Teacher, Infection and Immunity, graduate students
1993-1996	Teacher, Tumor Immunology, graduate students
1994-1996	Teacher, Advanced Cellular Immunology, graduate students
1994-1996	Teacher, Molecular Immunology and Immunogenetics, graduate students
1995-1996	Coordinator, Graduate Program Course on "Tumor Immunology"
1999	Teacher, AIDS Pathogenesis
2002-present	Roswell Park Cancer Institute
	Teacher, Advances in Immunology, graduate students

Trainees (Students and Postdoctoral Fellows)

Dariusz Kmieciak M.Sc.,
Martin Schmid M.Sc.,
Irena Kiszka, M.Sc.,
Zhien Wang M.D. Ph.D.,
Richard Wiaderkiewicz Ph.D.,
Jacek Bartkowiak Ph.D.,
Paul Jagodzinski Ph.D.,
Ilona Bednarek Ph.D.,
Marian Dmochowski M.D. Ph.D.,
Toshio Naito M.D. Ph.D.
Aleksandra Kowalczyk, M.Sc.,
Jolanta Bratosiewicz M.Sc.,
Andrew Wierzbicki M.Sc.,
Alexandra Kowalczyk, M.Sc.,
Elizabeth Hyjek M.D. Ph.D.,
Gregorz Kurzawski Ph.D.,
Richard Drozd Ph.D.,
Thomas Wasik Ph.D.,
Hiroshi Kaneko M.D. Ph.D.,
Jaroslaw Gzyl Ph.D.,
Elizabeth Bolesta, Ph.D.

Publications

PEER-REVIEWED PAPERS

Kozbor, D., Steinitiz, M., Klein, G., Koskimies, S., and Makela, O. Establishment of anti-TNP antibody-producing human lymphoid lines by preselection of hapten binding followed by EBV transformation. Scand. J. Immunol. 10: 187-194, 1979.

Kozbor, D. and Roder, J.C. Requirements for the establishment of high-titered human monoclonal antibodies against tetanus toxoid using the Epstein-Barr virus technique. J. Immunol. 127: 1275-1280, 1981.

Roder, J.C., Haliotis, T., Laing, L., **Kozbor, D.,** Rubin, P., Pross, M., Boxer, L., White, J., Fauci, A., Mostowski, H., and Matheson, D. The Chediak-Kigashi gene in human. III. Further studies on natural killer cell function in six patients. Immunology 46: 555-560, 1981.

Roder, J.C., Beaumont, T., Kerbel, R., Haliotis, T., and **Kozbor, D.** Selective natural killer resistance in a clone of YAC lymphoma cells. Proc. Natl. Acad. Sci. USA 78: 6396-6400, 1981.

Kozbor, D., Roder, J.C., Chang, T.H., Steplewski, Z., and Koprowski H. Human anti-tetanus toxoid monoclonal antibody secreted by EBV-transformed human B cells fused with a murine myeloma. Hybridoma 1(3): 323-328, 1982.

Kozbor, D., Lagarde, A.E., and Roder, J.C. Human hybridomas constructed with antigen specific EBV-transformed cell lines. Proc. Natl. Acad. Sci, USA 79: 6651-6655, 1982.

Kozbor, D., Dexter, D., and Roder, J.C. A comparative analysis of the phenotype characteristics of available fusion partners for the construction of human hybridomas. Hybridoma 2(1): 7-16, 1983.

Kozbor, D. and Roder, J.C. Comparison of a specific antibody response in humans induced by antigen (tetanus toxoid) or a polyclonal activator (EBV) in vitro. Inter. Arch. of Allergy and Applied Immunol. 72: 260-266, 1983.

Kozbor, D. and Roder, J.C. In vitro stimulated lymphocytes as a source of human hybridomas. Eur. J. Immunol. 14: 23-27, 1983.

Kozbor, D. and Croce, C.M. Amplification of the c-myc oncogene in a human breast carcinoma cell line. Cancer Research 44: 438-441, 1984.

Kozbor, D., Tripputi, P., Roder, J.C., and Croce, C.M. A human hybrid myeloma for production of human monoclonal antibodies. J. Immunol. 133(6): 3001-3005, 1984.

Cole, S.P.C., Campling, B.G., Louwman, I.H., **Kozbor, D.,** and Roder, J.C. A strategy for the production of human monoclonal antibodies reactive with lung tumor cell lines. Cancer Research 44: 2750-2753, 1984.

Cole, S.P.C., Campling, B.G. Atlaw, T., **Kozbor, D.,** and Roder, J.C. Human monoclonal antibodies. Mol. Cell. Biochem. 62: 109-120, 1984.

Atlaw, T., **Kozbor, D.,** and Roder, J.C. Human monoclonal antibodies against Mycobacterium leprae. Inf. Immun. 49: 104-110, 1985.

Kozbor, D., Abramow-Newerly, W., Tripputi, P., Cole, S.P., Weibel, J., Roder, J.C. and Croce, C.M. Specific immunoglobulin production and enhanced tumorigenicity following ascites growth of human hybridomas. J. Immunol. Meth. 81: 31-42, 1985.

Kozbor, D., Finan, J., Nowell, P.C., and Croce, C.M. The gene encoding the T4 antigen maps to human chromosome 12. J. Immunol., 136: 1141-1143, 1986.

Kozbor, D., Giallongo, A., Sierzega, M.E., Konopka, J.B., Witte, O.N., Showe, L.C., and Croce, C.M. Expression of a translocated c-abl gene in hybrids of mouse fibroblasts and chronic myelogenous leukemia cells. Nature, 319: 331-333, 1986.

Kozbor, D., Moretta, A., Messner, H.A., Moretta., L., and Croce, C.M. Tp44 molecules involved in antigen-independent T cell activation are expressed on human plasma cells. J. Immunol., 138: 4128-4132, 1987.

Kozbor, D., Burioni, R., ar-Rushdi, A., Zmijewski, C., and Croce, C.M. Expression of members of immunoglobulin gene family in somatic cell hybrids between human B-and T-cells. Proc. Natl. Acad. Sci. USA, 84: 4969-4973, 1987.

Kozbor, D., Trinchieri, G., Monos, D., Isobe, M., Russo, G., Haney, J., Zmijewski, C., and Croce, C.M. Human TCR $\gamma^{+}\delta^{+}$, $CD8^{+}$ T lymphocytes recognize tetanus toxoid in an MHC-restricted fashion. J. Exp. Med., 169: 1847-1851, 1989.

Kozbor, D., Casatella, M., Kagan, J., Finver, S., Faust, J., Trinchieri, G., and Croce, C.M. Expression and function of $\gamma\delta$ and $\alpha\ \beta$ T cell receptor heterodimers on human somatic T cell hybrid. J. Immunol.; 144: 3677-3683, 1990.

Gaulton G. N., Brass, L. F., **Kozbor, D.,** Pletcher C. H., and Hoxie J. A. Inhibition of T cell receptor-dependent phosphorylation of CD4 in human immunodeficiency virus type 1 infected cells. J. Biol. Chem. 267: 4102-4109, 1992.

Kozbor, D., Hyjek, E., Wiaderkiewicz, R., Wang Z., Wang, M., and Loh, E. Competitor mRNA fragments for quantitation of cytokine specific transcripts in cell lysates. Mol. Immunol. 30: 1-7, 1993.

Kozbor, D., Hyjek, E., Wiaderkiewicz, R., Kurzawski, G., and Lischner, H. W. Analysis of $\gamma\delta^+$ T cells in peripheral blood of children with perinatal human immunodeficiency virus (HIV) infection. J. Clin. Immunol. 13: 193-203, 1993.

Loh, E.Y., Wang, M., Bartkowiak, J., Wiaderkiewicz, R., Hyjek, E., Wang, Z. and **Kozbor, D.** Gene transfer studies of T cell receptor-$\gamma\delta$ recognition: Specificity for Staphylococcal Enterotoxin A is conveyed by Vγ9 alone. J. Immunol. 152: 3324-3332, 1994.

Jagodzinski, P.P., Wiaderkiewicz, R., Kurzawski, G., Kloczewiak, M., Nakashima, H., Hyjek, E., Yamamoto, N., Uryu, T., Kaneko, Y., Posner, M.R. and **Kozbor, D.** Mechanism of inhibitory effect of curdlan sulfate on HIV-1 infection in vitro. Virology 202:735-745, 1994.

Hyjek, E., Lischner, H. W., Hyslop, T. Bartkowiak, J., Kubin, M., Trinchieri, G., and **Kozbor, D.** Cytokine patterns during progression to AIDS in children with perinatal HIV-1 infection. J. Immunol., 155:4060-4071, 1995.

Jagodzinski, P., Wustner, J., Kmieciak, Wasik, T. J., Fertala, A., Sieron, A., Takahashi, M., Tsuji, T., Mimura, T., Fung, M.S., Gorny, M. K., Kloczewiak, M., Kaneko, Y. and **Kozbor, D.** Role of the V2, V3 and CD4-binding domains of gp120 in curdlan sulfate neutralization sensitivity of HIV-1 during infection of T lymphocytes. Virology, 226:217-227, 1996.

Hyjek, E. M., Bartkowiak, J., Drozdz, R., Wasik, T. J., Jasinski, M., Lischner, H. W. and **Kozbor, D.** Evidence for B cell-mediated activation of V$\delta1^+$ T lymphocytes during progression of HIV infection. J. Immunol. 158:464-474, 1997.

Wasik, T. J., Jagodzinski, P. P., Hyjek, E. M., Lischner, H. W., and **Kozbor, D.** Diminished HIV-specific CTL activity associated with enhanced type 2 responses to HIV-specific peptides during perinatal HIV infection. J. Immunol. 158:6029-6036, 1997.

Kmieciak, D., Jasinski, M., Teppler, H., Pientka, J., Hsu, S. H., Takahashi, H., Okumura, K., Kaneko, Y., and **Kozbor, D.** The effect of deletion of the V3 loop of gp120 on induction of cytotoxic T cell responses and HIV gp120-mediated pathogenesis. J. Immunol. 160:5676-5683, 1998.

Kmieciak, D., Bednarek, I., Takiguchi. M., Bratosiewicz, J., Wierzbicki, A., Wasik, T., Teppler, H., Pientka, J., Hsu, S. H., and **Kozbor, D.** The effect of epitope variation on the profile of CTL responses to the HIV envelope glycoprotein. International Immunol. 10:1789-1799, 1998.

Jagodzinski, P. P., Wustner, J., Kaneko, Y. and **Kozbor, D.** Enhanced binding of G3-4 monoclonal antibodies to the V2 domain of gp120 accompanies increasing infectivity of the THP-1 cells with M-tropic HIV-isolates. Viral Immunology, 12: 23-33, 1999.

Wasik, T. J., Lischner, H. W., Jasinski, M., Bratosiewicz, J., Whiteman, V. E., Rutstein, R., Starr, S. E., Douglas, S., Kaufman, D., Sison, A. V., and **Kozbor, D.** Protective role of nonlytic immune responses in vertical HIV transmission. J. Immunol. 162:4355-4364, 1999.

Wasik, T. J., Wierzbicki, A., Trinchieri, G., Whiteman, V. E., Lischner, H. W., and **Kozbor, D.** Early HIV-specific Th1 and CTL responses associated with slow disease progression in Infants with perinatal HIV infection. Eur. J. Immunol. 30:117-127, 2000.

Kaneko, H., Bednarek, I., Wierzbicki, A., Kiszka, I., Wasik, T. J., Kaneko, Y., and **Kozbor, D.** Oral DNA vaccination promotes mucosal and systemic cellular responses to HIV envelope glycoprotein. Virology 267:8-16, 2000.

Kmieciak, D., Bolesta E., Naito. T., Kaneko Y., and **Kozbor, D.** Enhancement of Cellular and Humoral Immune Responses to Human Immunodeficiency Virus Type 1 Gag and Pol by a GP-92 Fusion Protein Expressing Highly Immunogenic Gag p17/p24 and Pol p51 Antigens. 2001. J. Hum. Virol. 6:306-316, 2001.

Wierzbicki, A., Kiszka, I., Kaneko, H., Wasik, T. J., Kaneko, Y., and **Kozbor, D.** Immunization strategies to augment oral vaccination with DNA and viral vectors expressing HIV env glycoprotein. Vaccine. 3014:1-13, 2002.

Kiszka, I., Kmieciak, D., Gzyl, I., Naito, T., Kaneko, Y., and **Kozbor, D**. Effect of deletion of the V3 loop of envelope glycoprotein on cellular responses and protection against challenge with recombinant vaccinia virus expressing gp160 of primary HIV-1 isolates. J. Virology 76:4222-4232, 2002.

Gzyl, J., Bolesta, E., Wierzbicki, A., Kmieciak, D., Naito, T., Honda, M., Komuro, H., Kaneko, Y., and **Kozbor, D.** Effect of partial and complete variable loop deletions of the human immunodeficiency virus type 1 envelope glycoprotein on the breadth of gp160-specific responses. Virology. 318:493-506, 2004.

PAPERS, MONOGRAPHS, BOOK CHAPTERS, AUDIOTAPES, AND LETTERS TO THE EDITOR

Roder, J.C., Haliotis, T., and **Kozbor, D.** Genetic control of human NK function.
In: Human Cancer Immunology (Serrou and Rosenfeld, eds.). Elsevier/North Holland Biochemical Press, Vol. 4, pp. 170-188, 1982.

Kozbor, D. and Roder, J.C. Monoclonal antibodies produced by human lymphocytes. Immunology Today 4(3): 72-79, 1983.

Roder, J.C., Cole, S.P.C., Ataw, T., Campling, B.G., McGarry, R.C., and **Kozbor, D.** 1985 The EBV-Hybridoma technique. In: Human Hybridomas and Monoclonal Antibodies (E.G. Engleman, S. Foung, J. Larrick, and A. Raubitschek, eds.) Plenum Press, New York, pp. 55-71.

Kozbor, D. and Croce, C.M. Fusion partners for production of human monoclonal antibodies. In: Human Hybridomas and Monoclonal Antibodies (E.G. Engleman, S. Foung, J. Larrick, and A. Baubitschek, eds.) Plenem Press, New York, pp. 21-37, 2985.

Cole, S.P., **Kozbor, D.,** and Roder, J.C. 1985 Strategies for production of human monoclonal antibodies. In: Hybridomas in Biotechnology and Medicine (T.A. Springer, ed.), pp. 43-57.

Cole, S.P., **Kozbor D,** and Roder, J.C. 1985 The EBV-hybridoma technique and its application to human lung cancer. In: Monoclonal Antibodies and Cancer Therapy (R.A. Reisfel and S. Cell, eds.). UCLA Symposia on Molecular and Cell Biology, New Series, Vol. 27. A.R. Liss, Inc., New York.

Roder, J.C., Cole, S.P.C., and **Kozbor D.** 1986 Monoclonal antibodies produced by EBV-hybridoma technique. In: Methods in Enzymology, Immunological Techniques (S.P. Colowick and N.O. Kaplan, eds.). Academic Press, Inc., Orlando, FL. Vol. 121, pp. 140-167.

Kozbor, D., Sierzega, M.E., Roder, J.C., and Croce, C.M. 1986 Comparative phenotypic analysis of available hybridoma fusion partners. In: Methods in Enzymology, Immunological Techniques (S.P. Colowick and N.O. Kaplan, eds.). Academic Press, Inc., Orlando, FL. Vol. 121, pp. 120-140.

Campling, B.G., Cole, S.P.C., Atlaw, T., **Kozbor, D.,** and Roder, J.C. 1986 Practical aspects of human-human hybridomas. In: Monoclonal Antibodies: Hybridoma Techniques (L.B. Schook, ed.).

Burioni, R., Romano, J., Abramow-Newerly, W., Roder, J.C., **Kozbor, D.,** and Croce, C.M. 1986 Production of human monoclonal antibodies by the EBV-hybridoma technique. In: Immunoassay Technology (S.B. Pal, ed.).

Campling, B.G., Cole, S.P.C., Atlaw, T., Roder, J.C., and **Kozbor, D.** The EBV-hybridoma technique and its application. In: A.J. Strelkauskas (ed.), Human Hybridomas, pp. 3-22, 1987.

Kozbor, D., Burioni, R., Romano, J., and Croce, C.M. Development of the human hybridoma technology. In: Cellular, Molecular and Genetic Approaches to Immunodiagnosis and Immunotherapy, pp. 59-71. University of Tokyo Press, 1987.

Kozbor, D. and Croce, C.M. Human Hybridoma. Fusion partners. In: A.H. Bartal and Y. Hirshaut (eds.), Methods of Hybridoma Formation, pp. 273-298. Clifton, New Jersey: Humana Press, Inc. 1987.

Kozbor, D. In vitro-stimulated lymphocytes as a source of human monoclonal antibodies generated by the EBV-hybridoma technique. pp. 162-169. Elsevier Science Publishers, B.V., 1987.

Kozbor, D. In vitro stimulated lymphocytes as a source of human monoclonal antibodies generated by the EBV-hybridoma technique. Progress in Biotechnology, 5: 193-208, 1988. In vitro Immunization in Hybridoma Technology.

Burioni, R., Romano J., Abramow-Newerly,W., Roder, J.C., Croce, C., and **Kozbor, D.** The Epstein-Barr virus-hybridoma technique for production of human monoclonal antibodies. Reviews on Immunoassay Technology, 2:51-72, 1988.

Kozbor, D. The γδ T cell receptor specific for tetanus toxoid. In: Cellular Immunity and the Immunotherapy of Cancer, UCLA Symposia on Molecular and Cellular Biology, New Series, (M.T. Lotze and O.J. Finn) Wiley-Liss, Inc., New York, NY. pp 41-44, 1990.

APPENDIX

Appealed Claims 14, 15, 19, and 21-36 read as follows:

Claim 14. A method for preparing a vaccine against human immunodeficiency virus (HIV) comprising:

(a) introducing into a vector DNA or liposome a nucleic acid encoding an envelope glycoprotein of HIV, wherein said envelope glycoprotein comprises a deletion of the third variable loop (V3); and

(b) mixing said vector DNA or liposome with a suitable adjuvant.

Claim 15. The method of Claim 14, wherein said nucleic acid is introduced into antigen presenting cells (APCs) and said APCs are mixed with adjuvant.

Claim 19. A vaccine for inducing cellular immunity against HIV comprising:

(a) cells expressing on their surfaces an envelope glycoprotein of HIV, wherein said envelope glycoprotein comprises a deletion of the third variable loop (V3); and

(b) an adjuvant.

Claim 21. The method of Claim 14, wherein said deletion of the third variable loop (V3) comprises deletion of amino acids 297 to 329 in said variable loop.

Claim 22. The method of Claim 14, wherein said human immunodeficiency virus is human immunodeficiency virus 1 IIIB.

Claim 23. The vaccine of Claim 19, wherein said cells are antigen presenting cells.

Claim 24. The vaccine of Claim 19, wherein said deletion of the third variable loop (V3) comprises deletion of amino acids 297 to 329 in said variable loop.

Claim 25. The vaccine of Claim 19, wherein said human immunodeficiency virus is human immunodeficiency virus 1 IIIB.

Claim 26. A method for preparing a vaccine against human immunodeficiency virus (HIV) comprising:

(a) introducing into a vector DNA or liposome, a nucleic acid encoding an envelope glycoprotein of HIV 1-IIIB, wherein said envelope glycoprotein comprises a deletion of amino acids 297 to 329 in the third variable loop (V3); and

(b) mixing said vector DNA or liposome with a suitable adjuvant.

Claim 27. A vaccine for inducing cellular immunity against HIV comprising:

(a) antigen presenting cells expressing on their surfaces an envelope glycoprotein of HIV 1 IIIB, wherein said envelope glycoprotein comprises a deletion of amino acids 297 to 329 in the third variable loop (V3); and

(b) an adjuvant.

Claim 28. A method of preparing a composition for stimulating CTL activity against human immunodeficiency virus, comprising

(a) introducing into a vector DNA or liposome a nucleic acid encoding an envelope glycoprotein of HIV, wherein said envelope glycoprotein comprises a deletion of the third variable loop (V3); and

(b) mixing said vector DNA or liposome with a suitable adjuvant.

Claim 29. The method of Claim 28, wherein said nucleic acid is introduced into antigen presenting cells (APCs) and said APCs are mixed with adjuvant.

Claim 30. The method of Claim 28, wherein said deletion of the third variable loop (V3) comprises deletion of amino acids 297 to 329 in said variable loop.

Claim 31. The method of Claim 28, wherein said human immunodeficiency virus is human immunodeficiency virus 1 IIIB.

Claim 32. A method of stimulating a CTL response in a patient, comprising administering the composition prepared according to the method of Claim 28 to the patient in an amount sufficient to stimulate a CTL response.

Claim 33. A method of stimulating a CTL response in a patient, comprising administering the composition prepared according to the method of Claim 29 to the patient in an amount sufficient to stimulate a CTL response.

Claim 34. A method of stimulating a CTL response in a patient, comprising administering the composition prepared according to the method of Claim 30 to the patient in an amount sufficient to stimulate a CTL response.

Claim 35. A method of stimulating a CTL response in a patient, comprising administering the composition prepared according to the method of Claim 31 to the patient in an amount sufficient to stimulate a CTL response.

Claim 36. A method of stimulating a CTL response in a patient, comprising administering the composition prepared according to the method of Claim 32 to the patient in an amount sufficient to stimulate a CTL response.

AIDS RESEARCH AND HUMAN RETROVIRUSES
Volume 10, Number 9, 1994
Mary Ann Liebert, Inc., Publishers

Viral Variability and Serum Antibody Response in a Laboratory Worker Infected with HIV Type 1 (HTLV Type IIIB)

MARVIN S. REITZ, JR.,[1] LEOTA HALL,[1] MARJORIE ROBERT-GUROFF,[1] JAMES LAUTENBERGER,[2] BEATRICE M. HAHN,[3] GEORGE M. SHAW,[3] LILLY I. KONG,[3,4] STANLEY H. WEISS,[5,6] DAVE WATERS,[7] ROBERT C. GALLO,[1] and WILLIAM BLATTNER[5]

ABSTRACT

Molecular clones of HIV-1 were obtained from isolates cultured from peripheral blood mononuclear cells (PBMCs) and directly from uncultured PBMCs from a laboratory worker accidentally infected with the HIV-1 laboratory strain, HIV-1(HTLV-IIIB). Envelope sequences corresponding to the first 752 amino acids of HIV-1(HTLV-IIIB) clone BH10 were obtained from clones of cultured virus and sequenced. Three *env* clones obtained shortly after infection differed among themselves at only seven nucleotide positions, resulting in one amino acid substitution and one frameshift mutation. These envelope sequences were as similar to the envelope sequences of various IIIB clones as the latter were to each other. *env* divergence increased over the course of infection. However, the overall diversity in *env* clones obtained two or more years after infection was still comparable to that among IIIB *env* clones from the original IIIB culture. Multiple clones of partial *env* gene sequences containing the V3 loop were also obtained directly from uncultured PBMCs by polymerase chain reaction amplification. The *env* sequences of these clones were generally similar to those of the cultured viruses. Within the V3 region, the earliest isolates retained the sequence of the HXB2 clone from IIIB. Clones obtained later showed a progressive divergence in V3. An A-to-T substitution within the GPGRAF sequence at the tip of the V3 loop was observed within 1 year after infection, and this mutation predominated in all subsequent isolates. Antibodies against the V3 loops of IIIB and divergent 1987 and 1990 LW isolates appeared simultaneously in laboratory worker serum and persisted with no significant differences in titer. Furthermore, neutralization studies with autologous sequential sera suggested selection for the A-to-T change in V3 was not due to V3-directed antibodies. These results demonstrate a surprising homogeneity among *env* sequences of HIV-1 from an infected laboratory worker, perhaps because the initial infection originated from a relatively homogeneous population of tissue culture-adapted virus.

INTRODUCTION

ACCIDENTAL INFECTION of a laboratory worker (LW) with HIV-1 has been previously reported.[1] LW was involved in high-volume production of concentrated HIV-1 (human T cell leukemia/lymphotropic virus [HTLV]-IIIB). Although no obvious exposure incident could be documented and no symptoms of an acute HIV-1 seroconversion syndrome were apparent, seropositivity was noted during routine screening, and virus was subsequently isolated. Restriction endonuclease analysis by Southern blotting showed that the virus was closely related to HIV-1(HTLV-IIIB), indicating that transmission most likely

[1]Laboratory of Tumor Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892.
[2]National Cancer Institute, Frederick Cancer Research and Development Center, Frederick, Maryland 21701.
[3]Departments of Medicine and Microbiology, University of Alabama at Birmingham, Birmingham, Alabama 35294.
[4]Present address: Hillcrest Biologicals, 10703 Progress Way, Cypress, California 90630.
[5]Environmental Epidemiology Branch, National Cancer Institute, National Institutes of Health, Rockville, Maryland 20892.
[6]Present address: University of Medicine and Dentistry, New Jersey Medical School, 30 Bergen Street, ADMC 1614, Newark, New Jersey 07107.
[7]NCI/FCRDC, Program Resources, Inc./DynCorp, AIDS Vaccine Program, Human Retroviruses Section, Frederick, Maryland 21702.

occurred in the laboratory by an exposure that was not apparent. Since then, LW has not developed any signs or symptoms of disease except for a progressive loss in CD4+ T cells.

Genetic diversity, particularly in the *env* gene, is one of the notable features of HIV-1,[2-8] and is a serious problem in vaccine development. Not only is there considerable variability between isolates from different individuals, but there is substantial microheterogeneity within viral populations from single individuals. Such populations, which have been called "quasispecies" and "swarms," can differ substantially in selected regions of the *env* gene.[9-14] This diversity can be small immediately after infection,[15-17] but rapidly increases. Balfe *et al.*[9] report a mean nucleotide distance within single individuals of 4.2% (range, 3.2–6.1%) in the third hypervariable region (V3) and 5.5% (range, 3.8–9.2%) in V4 and V5 5 years after infection. The genetic makeup of the virus population as a whole also shifts over time, possibly owing to pressures from the host immune system. Immune selection pressure has been demonstrated by antibodies in natural antisera *in vitro*, resulting in neutralization-resistant mutants.[18] A single amino acid substitution in the envelope protein has been shown to be sufficient to generate escape mutants.[19,20]

In transmission from one individual to another, infection may involve a great variety of genotypes, yet, as indicated by the low initial diversity in the recipient, only a few may make up the majority of the initial population.[15] Transmission and growth of HIV-1 in culture appears to select subgroups within populations, resulting in a more limited degree of intrastrain variability.[21] LW was infected with a cultured virus with limited variability and for which considerable sequence information exists. This should simplify attempts to understand the genetic drift of virus populations over time in this individual. We have therefore sequenced 10 nearly complete *env* genes from 4 isolates obtained over a 5-year period from this individual and attempted to correlate this with the immune response. We have also obtained partial *env* sequences directly from peripheral blood cells by polymerase chain reaction (PCR) amplification.

MATERIALS AND METHODS

Immunological assays

The assay for neutralization of cell-free HIV-1 has been described in detail.[22] The peptide-binding assay has also been previously described.[23] Peroxidase conjugates of goat anti-human IgG or goat anti-mouse IgG were used as appropriate. Synthetic peptides were obtained from Multiple Peptide Systems, Inc. (Richmond, CA) and were used as 80% pure preparations. The amino acid sequences of the peptides used were as follows:

IIIB (BH10)	NNTRKSIRIQRGPGRAFVTIGKIG(C)
LW87	NNTRKRIRIQRGPGRTFVTIGKIG(C)
LW90	NNTGKRIRIQRGPGRTFVTIGKIG(C)

The carboxy-terminal cysteines were added to facilitate protein conjugation for eventual antibody production.

Virus isolation and sequence analyses

HIV-1 isolates were established and grown at various times from September 1985 to February 1990 in three independent laboratories from the peripheral blood mononuclear cells (PBMCs) of LW, a laboratory worker, accidentally infected with HIV-1(HTLV-IIIB). The PBMCs from LW were cocultured with uninfected normal PBMCs or with CEM or H9 cells as indicated. Reverse transcriptase (RT) assays were performed in 96-well microplates. Into each well was pipetted 20 μl of reaction mix, which consisted of 62.5 mM Tris-hydrochloride (pH 8.3), 0.25% Nonidet P-40 (NP-40), 125 mM NaCl, 7.5 mM $MgCl_2$, 12.5 mM dithiothreitol, oligo$(dT)_{12-18}$ (5 μg/ml), poly(rA) (50 μg/ml), and [^{3}H]ribosylthymine 5′-triphosphate ([^{3}H]TTP; 156 μCi/ml), followed by 6 μl of sample (equivalent to 50 μl of clarified culture fluid) derived by polyethylene glycol (PEG) precipitation. After 1 hr at 37°C the reaction was stopped, trichloroacetic acid (TCA) precipitated, and harvested onto glass filter mats and washed. The filter was dried and the samples punched into vials and counted.

Southern blot hybridization patterns of all LW isolates with seven different restriction endonucleases were virtually indistinguishable from each other and from that of one of the components of HIV-1(HTLV-IIIB). Two complete *env* genes were obtained from λ phage clones of HIV-1 from each of three independent PBMC isolates (LW85, LW87, and LW88). The number designates the year of draw of the blood samples from which the respective isolates were obtained. LW isolates lack in the central portion of the genome the *Sst*I site that is found in some HIV-1(HTLV-IIIB) genotypes.[1] Consequently, 8.9-kb *Sst*I fragments of LW proviral DNA were purified from infected PBMC cultures, cloned into the *Sst*I site of phage λgtWES.λB, and screened by standard methods.[24] Two clones were selected for sequence analyses from each of three libraries (clones LW85-1, 85-2, 87-1, 87-2, 88-1, and 88-2). These λ clones were digested with *Bam*HI and *Sal*I and the 2.7-kb fragment containing the *env* gene was purified and substituted for the homologous region of the infectious HIV-1 plasmid pHXB2gpt.[25]

A complete *env* gene sequence was obtained from an infectious molecular clone, LW12.3, representing a 1987 LW isolate.[26] To obtain clone LW12.3, H9 cells were infected with virus from macrophages that had previously been infected by cocultivation with a 1987 sample of PBMCs from LW, then biologically cloned by limiting dilution. High molecular weight DNA was purified from the biological clones by standard methods, digested with *Xba*I (which does not cut within HIV-1[IIIB] DNA), and ligated into the *Xba*I site of λ phage DASH II (Stratagene, La Jolla, CA). Two positive clones were identified from 10^6 recombinant phage and found to be identical by restriction endonuclease mapping. One of the two, LW12.3, was subcloned into an *Xba*I site that had been inserted into plasmid sP65gpt[25] with oligonucleotide linkers. All sequences were obtained from double-stranded plasmid DNA using the Sequenase II kit (United States Biochemical, Cleveland, OH) by the method of Sanger,[27] using primers based on the HXB2 sequence.[28]

The *env* gene sequences LW85-3, and LW90-1 and LW90-2, were obtained and assembled from three separate overlapping fragments of the *env* gene cloned from PCR amplification products. The target DNA was purified from cultured CEM cells cocultivated with PBMCs drawn on September 4, 1985 (9/4/85) and February 1, 1990 (2/1/90), respectively. Three sets of primers were used to amplify complete *env* genes from these isolates. These were as follows:

1A: 5′ ACGTTCTAGACATAGTATTCATA-GAATATAGGA 3′
1B: 5′ ACGTTCTAGACAGATCTAAATACTACCTCT 3′
3A: 5′ ACGTTCTAGAGGCAGTCTAGCAGAAG 3′
3B: 5′ ACGTTCTAGACCAGACTGTGAGTTGCAACAG 3′
4A: 5′ ACGTGAATTCGTTAGTGCAGCAGCAGAACA 3′
4B: 5′ ACGTAGATCTTACAGTAGGCCATCCAATCA

Primer pairs 1 and 3 contained *Xba*I sites (underlined) built into the primers, and amplified regions of the *env* gene corresponding to nucleotides 5733–6604 and 6604–7493, using the numbering system of Ratner *et al.*[29] Primer pair 4 contained an *Eco*RI site and a *Bgl*II site and amplified nucleotides 7462–8401. Amplification was for 30 cycles, consisting of denaturation (94°C, 1 min), annealing (55°C, 1 min), and extension (72°C, 2 min). After purification and cleavage by the appropriate enzyme(s), PCR fragments from the first two primers were cloned into the *Xba*I site of pBluescript KS (Stratagene) and from the last pair into the *Eco*RI and *Bam*HI sites of the same plasmid. Plasmid sequencing was by the method of Sanger.[27]

Polymerase chain reaction amplification and sequence analyses of the V3 region of the *env* gene from uncultured blood PBMCs was performed on samples drawn at five different times (2/4/86, 9/16/86, 3/19/87, 11/30/87, and 1/6/88). DNA was extracted in physically separated laboratories, using stringent precautions to avoid PCR carryover as well as sample contamination. Two oligonucleotide primers designed to amplify the V3 and adjacent envelope sequences were synthesized (Research Genetics, Huntsville, AL) and purified according to manufacturer recommendations. The sequences of the primers were as follows:

I: 5′ GTGTCAACTCAACTGCAGTTAAATGGCAGT 3′
II: 5′ TTGTTAACAGGATCCCTGTAATATTTGATG 3′

The primers contained *Pst*I and *Bam*HI restriction enzyme cleavage sites (underlined) to facilitate subsequent cloning into M13 phage and sequence analyses. The PCR reactions were carried out in a volume of 100 μl containing 4–8 μg of total PBMC genomic DNA, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin, 200 μM dXTPs, 20 pmol of each primer, and 2.5 U of *Taq* polymerase. Samples were overlaid with 100 μl of mineral oil to avoid evaporation, then subjected to 45 amplification cycles consisting of a denaturing step (94∞C, 1 min 30 sec), primer annealing step (50°C, 1 min 30 sec), and a primer extension step (72°C, 2 min 15 sec). Negative and positive control DNAs were amplified in parallel with the PBMC samples. The PCR products were gel purified, cleaved with *Pst*I and *Bam*HI, and cloned into M13mp19. Sixteen individual clones (designated by the prefix "Q"), representing the 5 different blood samples, were subsequently sequenced by the dideoxynucleotide chain termination method.[27]

The samples from which all of the different clones were obtained are summarized schematically in Fig. 1. Comparisons and alignments were made using the CLUSTAL program in the PC/Gene software package (Intelligenetics, Mountain View, CA) for amino acid sequences and the PILEUP program from the GCG package,[30] and the DNADIST, NEIGHBOR, SEQBOOT, and CONSENSE programs from the PHYLIP package

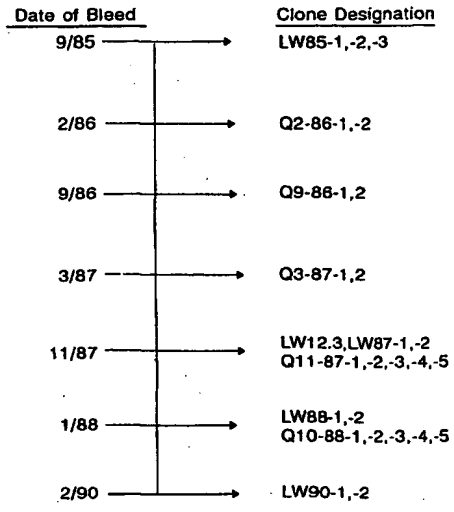


FIG. 1. *Left*: Dates on which blood samples were taken from LW and used to isolate virus or used directly for PCR. *Right*: *env* clones derived from LW isolates. Sequences prefixed with LW were obtained from cultured virus isolates, whereas those prefixed with Q were obtained by direct PCR of uncultured PBMCs.

(version 3.5c; J. Felsenstein, University of Washington, Seattle, Washington) were used for nucleotide sequences. Variability by amino acid residue was plotted using the method of Wu and Kabat.[31] Pairwise nucleotide comparisons were performed with the NALIGN program in PC/Gene, using the method of Myers and Miller.[32] The open gap cost and unit gap cost were both set at 10. DNA sequences from four individual clones of HIV-1(HTLV-IIIB) have been reported previously[28,29] and include BH10, BH8, HXB2, and HXB3.

RESULTS

Virus isolates from peripheral blood mononuclear cells of LW

Virus isolates were obtained by cocultivation of uninfected, phytohemagglutinin (PHA)-activated PBMCs with PBMCs from blood drawn from LW, who was apparently infected with the HIV-1 laboratory strain (HTLV-IIIB), at different times after the first detected seroconversion. As shown in Table 1, virus was readily grown in activated human PBMCs from the earliest blood sample (September 1985) to the latest (November 1991). As judged by the number of days required for RT activity to exceed given levels, which ranged from 4 to 45 days, there was no systematic overall increase in viral burden over this 6-year period. In addition to the efficient growth of these isolates on human peripheral blood lymphocyte (PBLs) they exhibited a characteristic IIIB cytopathic appearance. Many swollen bizarre-shaped cells associated with low levels of syncytia were observed, with most cell death attributed to the granulation of individual cells. All of the LW isolates obtained from 9/4/85 through 5/7/92 could be readily transmitted to T cell lines such as CEM or H9 using cell-free culture supernatant from infected PBL cultures.

TABLE 1. ISOLATION OF VIRUS FROM LW[a]

Blood sample[b] (month/year)	Growth in PBLs[c]: Time (days) in culture for RT t equal: 10^4 cpm/50 μl	10^5 cpm/50 μl
9/85	17	20
4/87	15	56
9/87	14	14
11/87	4	10
5/88	45	45
2/89	15	19
10/89	13	16
2/90	15	15
7/90	9	20
3/91	11	14
11/91	12	12

[a]Virus was isolated from peripheral blood cells by cocultivation with primary activated normal PBLs, as described in Materials and Methods.

[b]Date indicates the month and year that blood was drawn for virus isolation.

[c]Reverse transcriptase was assayed as described in Materials and Methods.

Comparison of envelope sequences of HIV-1 isolated from LW

Three molecular clones of virus isolates were obtained from blood drawn in September 1985, three from blood drawn in November 1987, two from blood drawn in January 1988, and two from blood drawn in February 1990. The DNA sequences of nearly the entire *env* genes, corresponding to the first 752 amino acids of the IIIB clone BH10, were determined. The 10 LW *env* sequences were compared with 16 representative HIV-1 sequences. A distance matrix representing the Jukes–Cantor distances was used to generate a phylogenetic tree (Fig. 2) by the neighbor-joining method. The robustness of the tree was analyzed by 100 bootstrap resamplings of the original data. There is close agreement between the topology of the neighbor-joining tree derived from the original data set and the consensus tree derived from the bootstrap replicas. In all 100 replications, the LW sequences form a monophyletic cluster within the HTLV-III/LAV group (node B). These sequences along with the laboratory worker sequences in turn formed a monophyletic cluster in the trees generated from all 100 replications (node D). Additionally, the North American/European group (corresponding to subgroup B of Myers *et al.*[33]) were clustered together in 100% of the replications (node E). Other nodes had lower values, indicating a lower degree of certainty in estimating the precise topology of the sequences within the North American/European group and within the HTLV-III/LAV cluster.

This analysis confirms that LW was infected by a virus from the HTLV-III/LAV group. Because all sequences in that group are derived from material isolated from one original patient early in the course of the epidemic, these data strongly support the hypothesis that LW was infected with the cultured virus that he/she was working with. With the exception of sequence LW12.3, sequences from isolates derived from the same blood samples (each drawn in a different year) clustered together. This suggests that at different times there was a single predominant form, although this interpretation is not certain because of the small number of genomes sequenced. The failure of the LW12.3 sequence to cluster with the other sequences derived from the 1987 PBMCs may be simply because its isolation in macrophages and subsequent passage history resulted in the isolation of a more variant virus.

As is evident from Table 2 and a phylogenetic tree of the aligned sequence (Fig. 2), the three 1985-derived envelope genes were more similar to each other than to any of four IIIB-derived clones, or than the IIIB clones are similar to each other. The close relationship of the two genomic clones to the PCR clone could not be due to cross-contamination, because the two isolates were grown, cloned, and sequenced at different times in different laboratories. The similarity of the LW85-3 sequence suggests that the PCR amplification did not result in a substantial error rate. The inferred amino acid sequence of the envelope proteins of the two clones from HIV-1(LW-1985) cultured in PBMCs were identical, whereas the PCR-generated clone from CEM cells differed by a single amino acid substitution (H for Y at position 250), the presence of a stop codon in place of tryptophan at position 672, and the insertion of a single G into the third base of the codon of residue 339, changing the amino acid sequence CNISRAKWN to CNISRAKWK and shifting the remaining reading frame out of phase. For the purposes of amino acid sequence comparisons, we removed the extra nucleotide and considered the stop codon as W. The resultant inferred translations were used for the Wu–Kabat plot in Fig. 3.

Blood drawn in November 1987 was also used for isolation of virus, as described previously.[1,34] Clones LW87-1 and -2 were obtained from viral isolates cultured in PBMCs, while LW12.3 is a molecular clone from virus that was cultured in peripheral blood monocyte/macrophages, then transmitted to H9 cells that were grown briefly and cloned by limiting dilution.[26] The envelope nucleotide sequences, as with the 1985 isolates, formed their own grouping on the phylogenetic tree (Fig. 2). Although the inferred envelope protein sequences of these 1987 LW isolates were more divergent than those from 1985 (Fig. 3), they still generally differed from each other less than they did from the IIIB clones or the IIIB clones did from each other. This is further shown by nucleotide difference scores generated from a series of pairwise comparison of the V3 and adjoining regions (Table 2) and a similar comparison of the V4 and adjoining regions (Table 3). These regions encompass the same regions analyzed by Balfe *et al.*[9] The overall envelope sequence of LW87-1 contained an insertion of five amino acids relative to the other two envelopes (Fig. 3); in addition, it differed by two amino acid substitutions from that of LW87-2 and by nine substitutions from that of LW12.3. The similarity of the LW12.3 sequence to those of LW87-1 and -2 is remarkable in view of their passage through different cell types in different laboratories.

Isolates were also obtained in 1988 and 1990 and grown in PBMCs and CEM cells, respectively. *env* sequences were obtained from λ clones (LW88-1 and -2) or from PCR-derived clones (LW90-1 and -2). These sequences showed a further divergence from the parental IIIB sequences (Fig. 2). In spite of

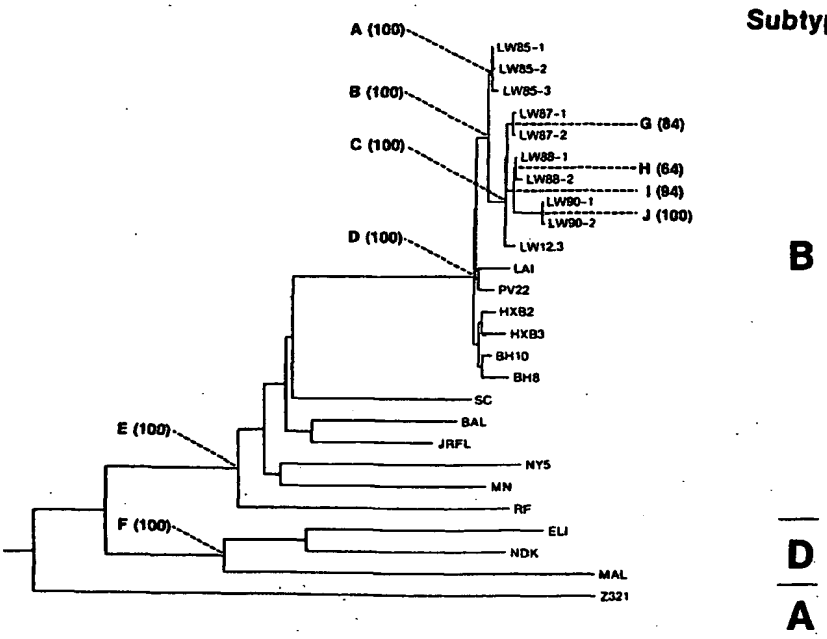


FIG. 2. Phylogenetic tree of *env* gene of HIV-1 isolates. The nucleotides sequences of partial *env* genes of the indicated HIV-1 and LW clones, corresponding to nucleotides 5802–8057 of the BH10 sequence of Ratner *et al.*,[29] were compared with each other and with the analogous *env* sequences from other representative HIV-1 isolates from the Human Retroviruses and AIDS Database.[33] The tree shown represents the neighbor-joining tree based on the Jukes–Cantor distances. Additionally, a consensus tree was derived from neighbor-joining trees constructed from 100 bootstrap replications of the data. The numbers to the right of the node labels represent the number of times the group consisting of the sequences to the right of the fork occur as a monophyletic cluster in this consensus tree. Prior to distance calculations, the sequences were aligned with the GCG PILEUP program.[30] The remaining analysis was accomplished by programs within the PHYLIP package. The subtypes shown are defined by Myers *et al.*[33] LW85-1 and -2, LW87-1 and -2, and LW88-1 and -2 are λ clones from PBMCs infected with blood obtained from LW in 1985, 1987, and 1988, respectively. LW85-3, and LW90-1 and -2, are from PCR clones of CEM cells infected from 1985 and 1990 blood samples, respectively. LW12.3 was obtained from a proviral λ clone of an isolate that was biologically cloned in infected H9 cells. This isolate originated from primary macrophages infected by virus from a 1987 peripheral blood sample.[1]

the two sets of sequences being derived from viruses isolated 2 years apart, their inferred envelope proteins were closely related, and formed their own group on the IIIB-LW branch of the dendrogram. Clone LW88-2 had an extra T inserted into a run of seven T residues in the DNA sequence at a position corresponding to the codons for amino acid residues 174–177. For the purpose of protein comparisons, the extra T was omitted from the translation. Aside from this change, the inferred envelope proteins of these four clones differed among themselves at only 17 positions. The V3 (Table 2) and V4 (Table 3) regions of these envelope sequences, as with the 1985 and 1987 isolates, were in general more closely related to each other than were those derived from the original IIIB clones. The data in Tables 2 and 3 do reveal, however, a gradual but distinct trend away from the envelope sequences of the original IIIB *in vitro* swarm and of the 1985 isolates, which is also evident from analysis of the whole *env* gene sequences shown in the phylogenetic tree in Fig. 2. The maximum deviation seen with the V3 regions are 6.1% nonidentity between the 1990 isolates and BH8. With the V4 sequences, the maximum nonidentity is 4.5%, between the 1990 isolates and BH8 (except for the 7.4% between LW87-1 and BH8, the alignment of which required the insertion of several gaps). Thus, the rate of divergence seen with these viruses is 1.2 and 0.9% per year in these two regions, respectively. As shown in the Wu–Kabat plot (Fig. 3),[31] a small cluster of differences among the LW90 and LW88 clones was evident at amino acid residues 575–625.

Changes in the V3 immunodominant loop of blood viral DNA

The V3 region, at residues 296–331 of the envelope protein, constitutes an immunodominant region that is the target of neutralizing antibodies.[35–37] We compared this region in clones obtained directly by PCR amplification from uncultured blood (designated with a prefix "Q") drawn at various times from LW. The first nucleotide was deleted from the codon for the carboxy-terminal C of the V3 loop in clones Q11-87-2 and Q11-87-4. These are indicated in Fig. 4 by an X. There was not a great difference between the clones from uncultured PBLs and cultured virus in the V3 region itself (Fig. 4) and in the adjacent regions, again suggesting an unusual lack of heterogeneity among the virus population in LW. Pairwise comparison of the nucleotide sequences of these clones, corresponding to nu-

TABLE 2. COMPARISON OF V3 REGIONS FROM LW ISOLATES

Isolate[b] \ Nucleotide differences[a]	HXB2	HXB3	BH8	BH10	LW85-1	LW85-2	LW85-3	LW87-1	LW87-2	LW12.3	LW88-1	LW88-2	LW90-1
HXB2													
HXB3	**0.013 (4)**												
BH8	**0.023 (7)**	**0.010 (3)**											
BH10	**0.023 (7)**	**0.037 (11)**	**0.033 (10)**										
LW85-1	0.020 (6)	0.034 (10)	0.037 (11)	0.010 (3)									
LW85-2	0.020 (6)	0.034 (10)	0.037 (11)	0.010 (3)	**0.000 (0)**								
LW85-3	0.020 (6)	0.034 (10)	0.037 (11)	0.010 (3)	**0.000 (0)**	**0.000 (0)**							
LW87-1	0.030 (9)	0.044 (13)	0.047 (14)	0.020 (6)	0.017 (5)	0.017 (5)	0.017 (5)						
LW87-2	0.034 (10)	0.047 (14)	0.050 (15)	0.024 (7)	0.020 (6)	0.020 (6)	0.020 (6)	**0.003 (1)**					
LW12.3	0.034 (10)	0.047 (14)	0.050 (15)	0.024 (7)	0.020 (6)	0.020 (6)	0.020 (6)	**0.010 (3)**	**0.007 (2)**				
LW88-1	0.037 (11)	0.050 (15)	0.054 (16)	0.027 (8)	0.024 (7)	0.024 (7)	0.024 (7)	0.007 (2)	0.003 (1)	0.010 (3)			
LW88-2	0.037 (11)	0.050 (15)	0.054 (16)	0.027 (8)	0.024 (7)	0.024 (7)	0.024 (7)	0.007 (2)	0.003 (1)	0.010 (3)	**0.000 (0)**		
LW90-1	0.044 (13)	0.057 (17)	0.061 (18)	0.034 (10)	0.030 (9)	0.030 (9)	0.030 (9)	0.013 (4)	0.010 (3)	0.017 (5)	**0.007 (2)**	**0.007 (2)**	
LW90-2	0.044 (13)	0.057 (17)	0.061 (18)	0.034 (10)	0.030 (9)	0.030 (9)	0.030 (9)	0.013 (4)	0.010 (3)	0.017 (5)	**0.007 (2)**	**0.007 (2)**	**0.000 (0)**

[a]Pairwise comparisons of V3 regions and adjacent sequences indicated above were made with the NALIGN program of PC/Gene, as described in Materials and Methods. Percent similarity scores obtained from the program were converted to difference scores by subtracting from 100 and multiplying by 0.01. The sequences compared encompassed the V3 region and correspond to nucleotides 6621 to 6917 (297 nt), using the numbering system of Ratner *et al.*[29] The numbers in parentheses indicate the number of substitutions between two sequences, and do not count gaps. Bold figures indicate nucleotide differences within clones from cultured virus, clones from 1985 LW isolates, clones from 1987 isolates, and clones from LW isolates obtained 1988–1990.

[b]Sequences were obtained from either plasmid clones of PCR-amplified DNA or λ phage clones of genomic DNA from cells infected *in vitro* as described in Materials and Methods.

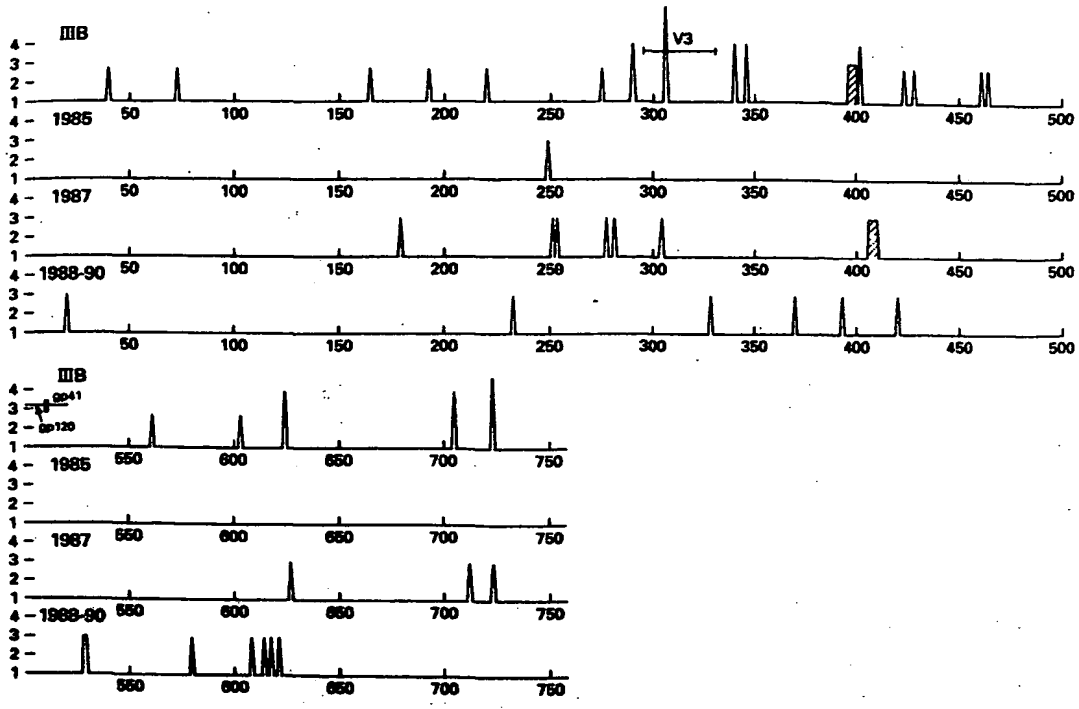


FIG. 3. Variability distribution at different times of isolation in HIV-1(HTLV-IIIB) and (LW) *env* genes. The inferred amino acid sequences of the envelope proteins of clones of HIV-1(HTLV-IIIB) and (LW) were aligned by the CLUSTAL program (PC/Gene, Intelligenetics, Inc.) and compared according to the method of Wu and Kabat.[31] Numbers below the lines refer to the inferred envelope amino acid position starting with the *env* gene initiation codon as residue 1. Peaks with diagonal hatching indicate differences due to deletions. LW85-3 and 88-2, as noted in Results, have single-nucleotide insertions that were removed for the purpose of protein alignments, and the former also has an in-frame stop codon that was considered as tryptophan for this comparison. IIIB indicates the comparison of BH8, BH10, HXB2, and HXB3 with each other; 1985 indicates the comparison of LW85-1, -2, and -3; 1987 indicates the comparison of LW87-1, -2, and LW12.3; and 1988–90 indicates the comparison of LW88-1, -2, and LW90-1.

cleotides 6575 to 7136 in the BH10 sequence,[29] showed that the sequences obtained by amplification of viral DNA in blood did not differ from each other (Table 4) more than from cultured virus (not shown) or more than the sequences of cultured viruses did from each other (Tables 2 and 3).

The sequences of the V3 loops from uncultured PBLs seemed to reflect a gradual drift from the sequence of the parental virus, and appeared to precede slightly the same changes in cultured viruses. The K-for-R substitution two residues downstream of the amino-terminal C of the V3 loop was prevalent in the PBL sequences from 1986 onward, but did not appear in the cultured virus sequences until 1988. Substitution of G for R eight residues downstream from the amino-terminal C was ubiquitous in PBLs by 1986, but R was still present in one of the sequences from virus cultured in 1987. A substitution in the V3 loop of GPGRTF for GPGRAF appeared to be ubiquitous from early 1986 onward. This change results in resistance to a neutralizing murine monoclonal antibody, M77.[38] All three clones from 1985 retained the HXB2 loop sequence, but all 22 clones obtained from 1986 to 1990 contained the T substitution. Within the cultured viruses, the 1990 isolates retained all the previous changes and, in addition, showed a change of QAHC to QVHC at the carboxy terminus of the loop. A clone obtained from a 1992 isolate retained this and previous changes and showed an additional change to RVHC.

Virus neutralization by autologous sera

Serum samples were taken from LW at different times, from September 1985 to July 1990, and tested for their ability to neutralize transmission of different viruses, including the infectious IIIB[28] molecular clone HXB2,[39] and the infectious clone LW12.3, obtained from a 1987 blood sample from LW.[26] LW12.3 has the GPGRTF sequence at the crown of the V3 loop, which is typical of all LW isolates obtained in 1986 or thereafter, and we wished to see whether this prevalence could be attributed to immunoselection against viruses with the IIIB GPGRAF V3 motif.

Sera drawn in September 1985, when seropositivity to HIV-1 was first detected, lacked measurable neutralizing activity, even at a dilution of 1:10. Serum drawn 6 months later was able to neutralize IIIB and LW12.3 at comparable titers (Table 5). Neutralization titers to both IIIB and LW12.3 continued to

TABLE 3. COMPARISON OF V4 REGIONS FROM LW ISOLATES

Isolate[b] / *Isolate*[b]	*Nucleotide differences*[a]												
	HXB2	*HXB3*	*BH8*	*BH10*	*LW85-1*	*LW85-2*	*LW85-3*	*LW87-1*	*LW87-2*	*LW12.3*	*LW88-1*	*LW88-2*	*LW90-1*
HXB2													
HXB3	**0.023 (6)**												
BH8	**0.008 (2)**	**0.025 (6)**											
BH10	**0.008 (2)**	**0.023 (6)**	**0.000 (0)**										
LW85-1	0.019 (5)	0.004 (1)	0.021 (5)	0.019 (5)									
LW85-2	0.019 (5)	0.004 (1)	0.021 (5)	0.019 (5)	**0.000 (0)**								
LW85-3	0.019 (5)	0.004 (1)	0.021 (5)	0.019 (5)	**0.000 (0)**	**0.000 (0)**							
LW87-1	0.021 (5)	0.004 (1)	0.074 (15)	0.021 (5)	0.000 (0)	0.000 (0)	0.000 (0)						
LW87-2	0.019 (5)	0.004 (1)	0.021 (5)	0.019 (5)	0.000 (0)	0.000 (0)	0.000 (0)	**0.000 (0)**					
LW12.3	0.019 (5)	0.004 (1)	0.021 (5)	0.019 (5)	0.000 (0)	0.000 (0)	0.000 (0)	**0.000 (0)**	**0.000 (0)**				
LW88-1	0.024 (6)	0.008 (2)	0.037 (9)	0.020 (5)	0.004 (1)	0.004 (1)	0.004 (1)	0.025 (3)	0.004 (1)	0.004 (1)			
LW88-2	0.024 (6)	0.008 (2)	0.037 (9)	0.020 (5)	0.004 (1)	0.004 (1)	0.004 (1)	0.025 (3)	0.004 (1)	0.004 (1)	**0.000 (0)**		
LW90-1	0.033 (8)	0.016 (4)	0.045 (11)	0.028 (7)	0.012 (3)	0.012 (3)	0.012 (3)	0.033 (5)	0.012 (3)	0.012 (3)	**0.008 (2)**	**0.008 (2)**	
LW90-2	0.033 (8)	0.016 (4)	0.045 (11)	0.028 (7)	0.012 (3)	0.012 (3)	0.012 (3)	0.033 (5)	0.012 (3)	0.012 (3)	**0.008 (2)**	**0.008 (2)**	**0.000 (0)**

[a]Pairwise comparisons of V4 regions and adjacent sequences indicated above were made and difference scores calculated as described in Table 2. The sequences compared encompassed the V4 region and correspond to nucleotides 6969 to 7226 (258 nt for all sequences except BH8 and LW87-1 [243 nt] and LW88-1 and –2 and LW90-1 and –2 [246 nt], using the numbering system of Ratner *et al.*[29] The numbers in parentheses indicate the number of substitutions between two sequences, and do not count gaps. Bold figures are for the comparison with the groupings described in Table 2.

[b]Sequences were obtained from either plasmid clones of PCR-amplified or DNA λ phage clones of genomic DNA from cells infected *in vitro* as described in Materials and Methods.

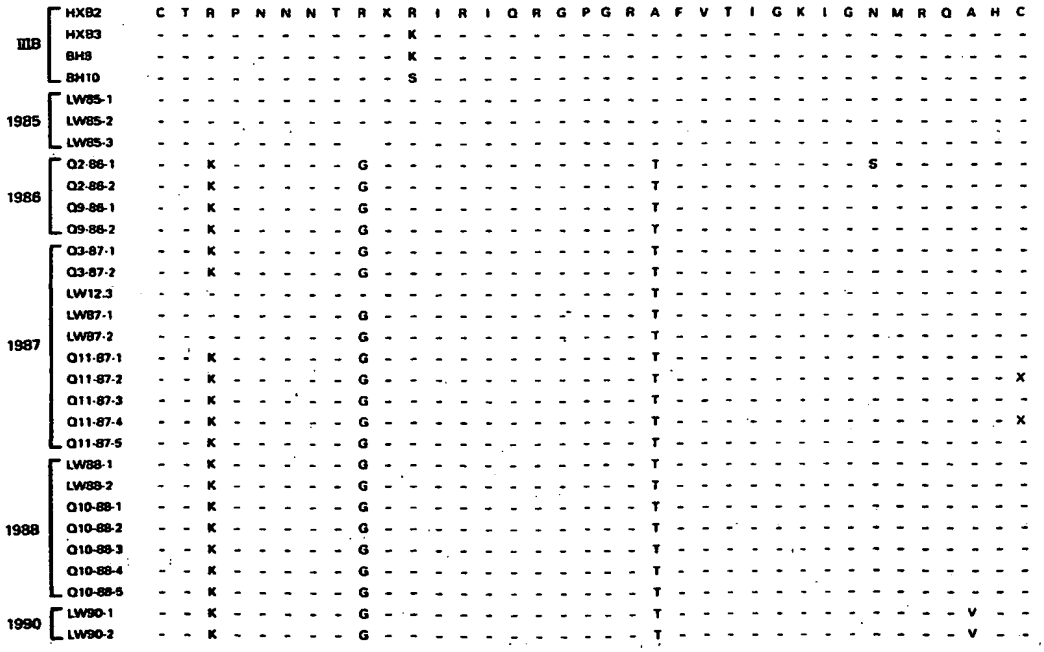


FIG. 4. Variation in V3 region with time. The amino acid sequences inferred from nucleotide sequences are shown for the indicated clones. The numbers on the left show the year that the blood was drawn from which viruses were isolated or that was used directly for PCR amplification. The LW sequences are from λ phage clones (LW85-1 and -2, LW87-1 and -2, LW12.3, LW88-1 and -2) or from PCR amplifications (LW85-3, LW90-1, and -2) using DNA from infected cell lines. The Q prefix indicates sequences that were obtained directly from uncultured blood cells by PCR amplification and molecular cloning. Q11-87-2 and -4, as noted in Results, had a deletion in the codon for the C-terminal cysteine of the V3 loop. For the comparison shown here, this is indicated by an X.

rise in parallel up to February 1990, the latest time point tested (Table 5). Neutralizing antibody titers against virus produced from the IIIB infectious clone pHXB2 were not as high as those against either IIIB itself, which contains a mixture of IIIB-related genomes,[28,29] or the LW12.3 clone.

Development of V3 loop-specific binding antibodies

Table 6, significant binding titers did not appear until early 1987, and they rose gradually thereafter. There were no apparent differences among the serial samples in recognition of the

TABLE 4. COMPARISON OF V3 SEQUENCES FROM BLOOD

V3 clone from blood[b]	*Nucleotide differences*[a]						
	Q2-86-1	*Q2-86-2*	*Q9-86-1*	*Q11-87-1*	*Q10-88-1*	*Q10-88-2*	*Q10-88-3*
Q2-86-1							
Q2-86-2	0.013 (7)						
Q9-86-1	0.013 (7)	0.000 (0)					
Q11-87-1	0.013 (7)	0.000 (0)	0.002 (1)				
Q10-88-1	0.013 (7)	0.000 (0)	0.002 (1)	0.000 (0)			
Q10-88-2	0.015 (8)	0.002 (1)	0.004 (2)	0.002 (1)	0.002 (1)		
Q10-88-3	0.013 (7)	0.000 (0)	0.002 (1)	0.002 (1)	0.000 (1)	0.002 (1)	

[a]Pairwise comparisons of V3 regions and adjacent sequences indicated above were made and difference scores calculated as described in Table 2. The sequences compared encompassed the V3 region and correspond to nucleotides 6575 to 7136 (550 nt for all sequences except Q2-86-1 and –2 [535 nt]), using the numbering system of Ratner *et al.*[29] Numbers in parentheses indicate the number of substitutions between two sequences, and do not count gaps.

[b]Sequences were obtained from plasmid clones of PCR-amplified DNA from primary blood cells taken from LW on the indicated dates as described in Materials and Methods.

TABLE 5. DEVELOPMENT OF NEUTRALIZING ANTIBODIES TO HIV-1 ISOLATES IN HIV-1(HTLV-IIIB)-INFECTED LABORATORY WORKER

Serum Date (month/year)	Neutralization titer[a] against: IIIB	HXB2	LW12.3
9/85	—	—	—
3/86	65	—	110
9/86	40	—	50
2/87	55	20	55
9/87	200	60	260
1/88	285	35	250
6/88	275	80	320
10/88	235	85	400
2/90	390	60	700
7/90	>440	115	ND

[a]Titer is expressed as the reciprocal of the serum dilution at which 60% of the target cells expressed viral p24, following normalization to infectivity in the presence of a standard neutralizing antibody-negative serum.

Abbreviation: ND, Not done.

BH10 loop, the LW87 loop with a single A-to-T substitution in the HXB2 sequence, or the LW90 loop with an additional R-to-G change in the region of the LW87 V3 sequence represented by the peptide.

DISCUSSION

We have characterized a series of HIV-1 isolates obtained from blood samples collected over a period of 5 years from a laboratory worker accidentally infected by cultured HIV-1(HTLV-IIIB). During this time, virus titer, as judged by the time required for detectable RT production following cocultivation of peripheral blood with uninfected target PBMCs, showed fluctuations but had no apparent trend. Virus from all time points grew readily in T cell lines CEM, H9, or both. Growth in established T cell lines is thought to result from the presence of more virulent ("rapid-high")[40] forms of HIV-1, and is usually more evident with virus from the latter stages rather than early in the course of infection.[40-42] LW, however, was infected by a virus that had been extensively adapted to permanent T cell lines. LW isolates (see below) showed little sequence divergence from IIIB during the 5-year period after infection, so it is perhaps not surprising that the LW isolates grow as well as the parental IIIB on permanent T cell lines.

TABLE 6. DEVELOPMENT OF BINDING ANTIBODIES TO HIV-1 V3 LOOPS IN SERA OF HIV-1(HTLV-IIIB)-INFECTED LABORATORY WORKER

Serum date (month/year)	Binding titer to V3 loops[a] BH10	LW87	LW90
9/85	<100	<100	<100
3/86	<100	<100	<140
9/86	<100	<100	<100
2/87	240	210	150
1/88	240	240	700
6/88	330	390	560
10/89	330	440	560
7/90	620	640	720

[a]Titer is expressed as the reciprocal of the lowest serum dilution at which the absorbance of the test serum equaled the absorbance of a standard negative serum diluted 1:100.

The analysis of the DNA sequences from the *env* genes of molecular clones of the LW isolates confirms that this laboratory worker was indeed infected by a virus from the HIV(HTLV-IIIB)/HIV-1(LAI) cluster. With the exception of LW12.3, sequences derived from the same blood samples (each drawn from a different year) clustered together. This may indicate that at any one time there is a single predominant form but over time this form drifts.

The *env* genes of three molecular clones of virus isolated in 1985, shortly after infection, showed little heterogeneity. Two clones obtained from λ phage libraries were identical, and a clone obtained by PCR amplification performed in a different laboratory from a separate culture differed from the λ clones by only three nucleotides. The *env* gene sequences of isolates obtained in 1987[1] were also compared. Two λ clones were obtained from infected normal PBMCs and a third was obtained from virus isolated and maintained in primary macrophage-monocytes, then passaged briefly in H9 cells and biologically cloned. Differences among these three 1987 clones were greater than among the 1985 isolates, although they were still comparable in degree to the differences observed among *env* sequences from the HIV-1(HTLV-IIIB)-derived clones and between sequences of the IIIB clones and the 1987 isolates. The failure of the LW12.3 sequence to cluster with the other 1987-derived sequences may be due to its initial selection in macrophages prior to molecular cloning and sequencing. *env* gene sequences taken from clones of isolates obtained in 1988 and 1990 were still remarkably similar to those of the original IIIB clones, as well as to each other, when compared with other intrapatient differences that have been reported.[9-14] As a comparative example, Balfe *et al.*[9] report intrapatient differences within a hemophiliac population of 3.2–6.1% for the V3 region obtained from blood DNA. The differences in V3 reported here at given time points range from 0 to 1.0% for cultured isolates and from 0 to 1.5% for viral DNA molecules in blood. Similarly, Balfe *et al.*[9] report intrapatient differences of from 1.0 to 9.2% in the V4 region, while among cultured viruses obtained from different time points we see differences of only 0–0.8%.

It was possible that this unusual homogeneity was due to selection of restricted genotypes in virus isolation and growth; it has been observed that culturing primary HIV-1 isolates can select for minor variants and minimize apparent intrapatient diversity.[11,14,21,43] This does not account for the low diversity seen in the present case, however, because the sequences of V3 and adjacent regions obtained directly from blood by PCR amplification are similar to the sequences obtained from cultured virus as well as to each other. A more likely contribution to the remarkable homogeneity seen in these LW isolates is that infection was by virus grown continuously in culture in T cell lines, which would be more homogeneous than virus populations within infected individuals.

The drift in the sequence of the V3 loop of the viral genomes appears to be significant and occurred more rapidly than that of the *env* gene as a whole. All nucleotide changes observed led to amino acid substitutions, and most substitutions were retained in subsequent isolates, resulting in a steady drift from the original V3 loops found shortly after seroconversion. The changes in the V3 loops of viral DNA from uncultured PBLs seemed to precede slightly the same changes from cells infected in culture. The drift in the V3 loop suggested that the variants observed might have originated because of selective pressure from neutralizing antibodies. We have previously observed[38] that the prevalent GPGRAF-to-GPRGTF change results in resistance to neutralization by a murine monoclonal antibody specific for the IIIB V3. To address this possibility, we tested serum samples taken from LW at various times for their ability to neutralize the original infecting virus, IIIB, as well as infectious clones derived from IIIB (HXB2, the V3 loop of which was identical with that of the earliest viral isolates from LW) and from a 1987 LW blood sample (LW12.3, which differs from HXB2 by substitution of T for A in the GPGRAF sequence at the crown of the IIIB V3 loop). Neutralizing antibodies against IIIB were detectable within 6 months of the first detected seroconversion, and neutralization titers continued to rise through the time period of the last sample tested, which was almost 5 years after seroconversion. Neutralizing antibodies to LW12.3 were also first detected 6 months after seroconversion at levels comparable to those seen against IIIB, and LW12.3 titers increased at the same rate as that seen with IIIB. This suggests that the rapid dominance of GPGRTF V3 loop variants is not due to selection *in vivo* by neutralizing antibodies directed against the IIIB V3 loop, although it does not rule out that selective pressure could be operating on other neutralizing epitopes or at the level of cellular immunity. We have previously shown that immunoselection *in vitro* with natural human antisera[18] can result in changes that do not involve the V3 loop,[19] and that seem to be conformational.[20] In view of previous reports of immune escape mutants arising soon after infection *in vivo*,[44] the failure to observe any evidence of *in vivo* immune selection here, at least up to 1987, was unexpected. Identification of the factors resulting in the observed genetic drift may provide insights into other important functional regions of the HIV-1 envelope proteins.

It is surprising that neutralizing antibodies against the IIIB molecular clone HXB2 were not observed in LW sera until 1 year after neutralizing antibodies first appeared against IIIB and LW12.3. HXB2 represents one of the genotypes from the IIIB culture, and might be expected to be neutralized at least as well as IIIB itself. Because the earliest viral clones obtained from LW had precisely the V3 loop sequence of HXB2, if the neutralizing antibodies were directed against V3, they should have been most effective against viruses with the HXB2 loop, including HXB2 itself. These data suggest that the initial neutralization is not directed against the V3 region, or at least not solely. This is borne out by the binding antibody data, where V3 antibodies were not detected until 1987 (Table 6). Alternatively, early neutralizing antibodies may be dependent on conformational epitopes formed in part by the V3 loop. This possibility is in keeping with the early type specificity observed in the neutralizing antibody response.

Two years after the first detection of neutralizing antibodies, the neutralization response broadened to include HIV-1(RF) (M. Robert-Guroff, unpublished results), a virus differing substantially from the IIIB-LW12.3 strains. This indicates the lack of broadly group-specific neutralizing antibodies until about 2.5 years after seroconversion. It is interesting that broadly neutralizing antibodies arise in the absence of detectable viral genomes that differ greatly from each other or from the infecting virus. This suggests that the broad neutralizing response is directed against conserved epitopes that are inherently less immunogenic than, for example, portions of the V3 loop. The conserved tip of the V3 loop (GPGR), however, may also contribute to the broadened response. Javaherian *et al.* have reported a generation of broadly neutralizing antibodies in response to a GPGRAF peptide.[45]

ACKNOWLEDGMENTS

This work was supported in part by grants from the National Institutes of Health and the U.S. Army Medical Research Acquisition Activity. G.M.S. is a PEW Scholar in the Biomedical Sciences.

REFERENCES

1. Weiss SH, Goedert JJ, Gartner S, Popovic M, Waters D, Markham P, di Marzo Veronese F, Gail MH, Barkley WE, Gibbons J, Gill FA, Leather M, Shaw GM, Gallo RC, and Blattner WA: Risk of human immunodeficiency virus (HIV-1) infection among laboratory workers. Science 1988;239:68–71.
2. Alizon M, Wain-Hobson S, Montagnier L, and Sonigo P: Genetic variability of the AIDS virus: Nucleotide sequence analysis of two isolates from African patients. Cell 1986;46:63–74.
3. Gurgo C, Guo HG, Franchini G, Aldovini A, Collalti E, Farrell K, Wong-Staal F, Gallo RC, and Reitz MS Jr: Envelope sequences of two new United States HIV-1 isolates. Virology 1988;164:531–536.
4. Hahn BH, Gonda MA, Shaw GM, Popovic M, Hoxie JA, Gallo RC, and Wong-Staal F: Genomic diversity of the acquired immune deficiency syndrome virus HTLV-III: Different viruses exhibit greatest divergence in their envelope genes. Proc Natl Acad Sci USA 1985;82:4813–4817.
5. Srinivasan A, Anand R, York D, Ranganathan P, Feorino P, Schochetman G, Curran J, Kalyanaraman VS, Luciw PA, and Sanchez-Pescador R: Molecular characterization of human immunodeficiency virus from Zaire: Nucleotide sequence analysis identifies conserved and variable domains in the envelope gene. Gene 1987;52:71–82.
6. Starcich BR, Hahn BH, Shaw GM, McNeely PD, Modrow S, Wolf H, Parks ES, Parks WP, Josephs SF, Gallo RC, and Wong-Staal F: Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS. Cell 1986;45:637–648.
7. Willey RL, Rutledge RA, Dias S, Folks T, Theodore T, Buckler CE, and Martin MA: Identification of conserved and divergent domains within the envelope gene of the acquired immunodeficiency syndrome retrovirus. Proc Natl Acad Sci USA 1986;83: 5038–5042.
8. Wong-Staal F, Shaw GM, Hahn BH, Salahuddin SZ, Popovic M, Markham P, Redfield R, and Gallo RC: Genomic diversity of human T-lymphotropic virus type III (HTLV-III). Science 1985;229: 759–762.
9. Balfe P, Simmonds P, Ludlam CA, Bishop JO, and Brown AJL:

Concurrent evolution of human immunodeficiency virus type 1 in patients infected from the same source: Rate of sequence change and low frequency of inactivating mutations. J Virol 1990; 64:6221–6233.
10. Fisher AG, Ensoli B, Looney D, Rose A, Gallo RC, Saag MS, Shaw GM, Hahn BH, and Wong-Staal F: Biologically diverse molecular variants within a single HIV-1 isolate. Nature (London) 1988;334:444–447.
11. Kusumi K, Conway B, Cunningham S, Berson A, Evans C, Iversen AKN, Colvin D, Gallo MV, Coutre S, Shpaer EG, Faulkner DV, DeRonde A, Volkman S, Williams C, Hirsch MS, and Mullins JI: Human immunodeficiency virus type 1 envelope gene structure and diversity in vivo and after cocultivation in vitro. J Virol 1992;66:875–885.
12. Saag MS, Hahn BH, Gibbons J, Li YX, Parks ES, Parks WP, and Shaw GM: Extensive variation of human immunodeficiency virus type-1 in vivo. Nature (London) 1988;334:440–444.
13. Simmonds P, Balfe P, Ludlam CA, Bishop JO, and Brown AJL: Analysis of sequence diversity in hypervariable regions of the external glycoprotein of human immunodeficiency virus type 1. J Virol 1990;64:5840–5850.
14. Vartanian JP, Meyerhans A, Åsjö B, and Wain-Hobson S: Selection, recombination, and G→A hypermutation of human immunodeficiency virus type 1 genomes. J Virol 1991;65: 1779–1788.
15. Cichutek K, Norley S, Linde R, Kreuz W, Gahr M, Lower J, von Wangenheim G, and Kurth R: Lack of HIV-1 V3 region sequence diversity in two hemophiliac patients infected with a putative biologic clone of HIV-1. AIDS 1991;5:1185–1187.
16. Zhang LQ, MacKenzie P, Cleland A, Holmes EC, Leigh Brown AJ, and Simmonds P: Selection for specific sequences in the external envelope protein of human immunodeficiency virus type 1 upon primary infection. J Virol 1993;67:3345–3356.
17. Zhu T, Mo H, Wang N, Nam DS, Cao Y, Koup RA, and Ho DD: Genotypic and phenotypic characterization of HIV-1 in patients with primary infection. Science 1993;261:1179–1181.
18. Robert-Guroff M, Reitz MS Jr, Robey WG, and Gallo RC: In vitro generation of an HTLV-III variant by neutralizing antibody. J Immunol 1986;137:3306–3309.
19. Reitz MS Jr, Wilson C, Naugle C, Gallo RC, and Robert-Guroff M: Generation of a neutralization-resistant variant of HIV-1 is due to selection for a point mutation in the envelope gene. Cell 1988;54:57–63.
20. Wilson C, Reitz MS Jr, Aldrich K, Klasse PJ, Blomberg J, Gallo RC, and Robert-Guroff M: The site of an immune-selected point mutation in the transmembrane protein of human immunodeficiency virus type 1 does not constitute the neutralization epitope. J Virol 1990;64:3240–3248.
21. Kusumi K, Conway B, Cunningham S, Berson A, Evans C, Iversen AKN, Colvin D, Gallo MV, Coutre S, Shpaer EG, Faulkner DV, DeRonde A, Volkman S, Williams C, Hirsch MS, and Mullins JI: Human immunodeficiency virus type 1 envelope gene structure and diversity in vivo and after cocultivation in vitro. J Virol 1992;66:875–885.
22. Robert-Guroff M: Neutralizing antibodies. In: *Techniques in HIV Research.* Aldovini A and Walker BD (eds.). Stockton Press, New York, 1990, pp. 179–185.
23. Robert-Guroff M, Aldrich K, Muldoon R, Stern TL, Bansal GP, Matthews TJ, Markham PD, Gallo RC, and Franchini G: Cross-neutralization of HIV-1, HIV-2, and SIV isolates. J Virol 1992;66:3602–3608.
24. Sambrook J, Fritsch EF, and Maniatis T: *Molecular Cloning.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989.
25. Feinberg MB, Jarrett RF, Aldovini A, Gallo RC, and Wong-Staal F: HTLV-III expression and production involve complex regulation at the levels of splicing and translation of viral RNA. Cell 1986;46:807–817.
26. Lori F, Hall L, Lusso P, Popovic M, Markham P, Franchini G, and Reitz MS Jr: Effect of reciprocal complementation of two defective human immunodeficiency virus type 1 (HIV-1) molecular clones on HIV-1 cell tropism and virulence. J Virol 1992;66:5553–5560.
27. Sanger F, Nicklen S, and Coulson AR: DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 1977;74: 5463–5467.
28. Ratner L, Fisher A, Jagodzinski LL, Mitsuya H, Liou RS, Gallo RC, and Wong-Staal F: Complete nucleotide sequences of functional clones of the AIDS virus. AIDS Res Hum Retroviruses 1987;3:57–69.
29. Ratner L, Haseltine W, Patarca R, Livak KJ, Starcich B, Josephs SF, Doran ER, Rafalski JA, Whitehorn EA, Baumeister K, Ivanoff L, Petteway SR, Pearson ML, Lautenberger JA, Papas TS, Ghrayeb J, Chang NT, Gallo RC, and Wong-Staal F: Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature (London) 1985;313:277–284.
30. Devereaux J: *Sequence Analysis Software Package Program Manual,* Version 7.0. Genetics Computer Group, Inc., Madison, Wisconsin, 1991, pp. 7:3–7:15.
31. Wu TT and Kabat EA: An analysis of the sequences of variable regions of Bence-Jones proteins and myeloma light chains and their implications for antibody complementarily. J Exp Med 1970;132:211–250.
32. Myers EW and Miller W: Optimal alignments in linear space. Comput Appl Biosci 1988;4:11–17.
33. Myers G, Berzofsky JA, Korber B, Smith RF, and Pavlakis G: *Human Retroviruses and AIDS 1991.* Los Alamos National Laboratory, Los Alamos, New Mexico, 1991.
34. Popovic M, Sarngadharan MG, Read E, and Gallo RC: Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. Science 1984;224:497–500.
35. Matsushita S, Robert-Guroff M, Rusche J, Koito A, Hattori T, Hoshino H, Javaherian K, Takatsuki K, and Putney S: Characterization of a human immunodeficiency virus neutralizing monoclonal antibody and mapping of the neutralizing epitope. J Virol 1988;62:2107–2114.
36. Rusche JR, Javaherian K, McDanal C, Petro J, Lynn DL, Grimaila R, Langlois A, Gallo RC, Arthur LO, Fischinger PJ, Bolognesi DP, Putney SD, and Matthews TJ: Antibodies that inhibit fusion of human immunodeficiency virus-infected cells bind a 24-amino acid sequence of the viral envelope, gp120 [published errata in Proc Natl Acad Sci USA 1988 Nov;85(22):8697 and Proc Natl Acad Sci USA 1989 Mar;86(5):1667]. Proc Natl Acad Sci USA 1988;85:3198–3202.
37. Skinner MA, Ting R, Langlois AJ, Weinhold KJ, Lyerly HK, Javaherian K, and Matthews TJ: Characteristics of a neutralizing monoclonal antibody to the HIV envelope glycoprotein. AIDS Res Hum Retroviruses 1988;4:187–197.
38. Di Marzo Veronese F, Reitz MS, Gupta G, Robert-Guroff M, Boyer-Thompson C, Louie A, Gallo RC, and Lusso P: Loss of a neutralizing epitope by a spontaneous point mutation in the V3 loop of HIV-1 isolated from an infected laboratory worker. J Biol Chem 1993;268:25894–25901.
39. Fisher AG, Collalti E, Ratner L, Gallo RC, and Wong-Staal F: A molecular clone of HTLV-III with biological activity. Nature (London) 1985;316:262–265.
40. Fenyo EM, Morfeldt-Manson L, Chiodi F, Lind B, von Gegerfelt A, Albert J, Olausson E, and Åsjö B: Distinct replicative and cytopathic characteristics of human immunodeficiency virus isolates. J Virol 1988;62:4414–4419.
41. Coombs RW, Collier AC, Allain JP, Nikora B, Leuther M, Gjesert

GF, and Corey L: Plasma viremia in human immunodeficiency virus infection. N Engl J Med 1989;321:1626–1631.

42. Ho DD, Moudgil T, and Alam M: Quantitation of human immunodeficiency virus type 1 in the blood of infected persons. N Engl J Med 1989;321:1621–1625.
43. Meyerhans A, Cheynier R, Albert J, Seth M, Kwok S, Sninsky J, Morfeldt-Manson L, Åsjö B, and Wain-Hobson S: Temporal fluctuations in HIV quasispecies in vivo are not reflected by sequential HIV isolations. Cell 1989;58:901–910.
44. Albert J, Abrahamsson B, Nagy K, Aurelius E, Gaines H, Nystrom G, and Fenyo E: Rapid development of isolate-specific neutralizing antibodies after primary HIV-1 infection and consequent emergence of virus variants which resist neutralization by autologous sera. AIDS 1990;4:107–112.
45. Javaherian K, Langlois AJ, LaRosa GJ, Profy AT, Bolognesi DP, Herlihy WC, Putney SD, and Matthews TJ: Broadly neutralizing antibodies elicited by the hypervariable neutralizing determinant of HIV-1. Science 1990;250:1590–1593.

Address reprint requests to:
Marvin S. Reitz
Laboratory of Tumor Cell Biology
National Cancer Institute
Building 37, Room 6A09
9000 Rockville Pike
Bethesda, Maryland 20892

GP160 sequences from HIV-1 IIIb variants

Line 1: HIV-1 IIIb (Leonard, et. al., Figure 1a)
Line 2: HIV-1 BH-10 isolate (accession no. P03375)
Line 3: HIV-1 HXB2 (accession no. P04578)
Line 4: HIV-1(HTLV IIIB, accession no. AAA76690)
Line 5: HIV-1 IIIb variant (Emini, et. al.,Figure 13-1)(partial)

```
                                        teklwvtvyy gvpvwkeatt tlfcasdaka
   1    mrvkekyqhl wrwgwrwgtm llgmlmicsa teklwvtvyy gvpvwkeatt tlfcasdaka
   1    mrvkekyqhl wrwgwrwgtm llgmlmicsa teklwvtvyy gvpvwkeatt tlfcasdaka
   1    mrvkekyqhl rrwgwrwgtm llgmlmicsa teklwvtvyy gvpvwkeatt tlfcasdaka
        ◄──────── Signal peptide ────────►

        ydtevhnvwa thacvptdpn pqevvlvnvt enfnmwkndm veqmhediis lwdqslkpcv
  61    ydtevhnvwa thacvptdpn pqevvlvnvt enfnmwkndm veqmhediis lwdqslkpcv
  61    ydtevhnvwa thacvptdpn pqevvlvnvt enfnmwkndm veqmhediis lwdqslkpcv
  61    ydtevhnvwa thacvptdpn pqevvlvnvt enfnmwkndm veqmhediis lwdqslkpcv

        kltplcvslk ctdlkndtnt nsssgrmime kgeikncsfn istsirgkvq keyaffykld
 121    kltplcvslk ctdlkndtnt nsssgrmime kgeikncsfn istsirgkvq keyaffykld
 121    kltplcvslk ctdlkndtnt nsssgrmime kgeikncsfn istsirgkvq keyaffykld
 121    kltplcvslk ctdlkndtnt nsssggmime kgeikncsfn istsirgkvq keyaffykhd

        iipidndtts ytltscntsv itqacpkvsf epipihycap agfailkcnn ktfngtgpct
 181    iipidndtts ytltscntsv itqacpkvsf epipihycap agfailkcnn ktfngtgpct
 181    iipidndtts yklltscntsv itqacpkvsf epipihycap agfailkcnn ktfngtgpct
 181    iipidndtts ytltscntsv itqacpkvsf epipihycap agfailkcnn ktfngtgpct
                                                          297 ──┐ Loop forming
                                                                │ Cys residues
        nvstvqcthg irpvvstqll lngslaeeev virsanftdn aktiivqlnq sveinctrpn
 241    nvstvqcthg irpvvstqll lngslaeeev virsanftdn aktiivqlnq sveinctrpn
 241    nvstvqcthg irpvvstqll lngslaeeev virsvnftdn aktiivqlnt sveinctrpn
 241    nvstvqcthg ikpvvstqll lngslaeeev virsanltdn vktiivqlnq sveinctrpn
                                                          q sveinctrpn
                                  329 ──┐                       │
        nntrksiriq rgpgrafvti gkignmrqah cnisrakwnn tlkqidsklr eqfgnnktii
 301    nntrksiriq rgpgrafvti gkignmrqah cnisrakwnn tlkqidsklr eqfgnnktii
 301    nntrkririq rgpgrafvti gkignmrqah cnisrakwnn tlkqiasklr eqfgnnktii
 301    nntrkririq rgpgrtfvti gkignmrqah cnisrakwnn tlkqiasklr eqygnnktii
        nntrksiriq rgpgrafvti gkignmrqah cnisrak

        fkqssggdpe ivthsfncgg effycnstql fnstwfnstw stegsnnteg sdtitlpcri
 361    fkqssggdpe ivthsfncgg effycnstql fnstwfnstw stkgsnnteg sdtitlpcri
 361    fkqssggdpe ivthsfncgg effycnstql fnstwfnstw stegsnnteg sdtitlpcri
 361    fkqssggdie ivthsfncgg effycnstql fnstwfnstw stegsnnteg sdtitlpcri

        kqfinmwqev gkamyappis gqircssnit glltrdggn nnneseifrp gggdmrdnwr
 421    kqiinmwqev gkamyappis gqircssnit glltrdggn snneseifrp gggdmrdnwr
 421    kqiinmwqkv gkamyappis gqircssnit glltrdggn snneseifrp gggdmrdnwr
 421    kqiinmwqev gkamyappis gqircssnit glltrdggn nnngseifrp gggdmrdnwr

        selykykvvk ieplgvaptk akrrvvqre
 481    selykykvvk ieplgvaptk akrrvvqrek ravgigalfl gflgaagstm gaasmtltvq
 481    selykykvvk ieplgvaptk akrrvvqrek ravgigalfl gflgaagstm gaasmtltvq
 481    selykykvvk ieplgvaptk akrrvvqrek ravgigalfl gflgaagstm gaasmtltvq

 541    arqllsgivq qqnnllraie aqqhllqltv wgikqlqari laverylkdq qllgiwgcsg
 541    arqllsgivq qqnnllraie aqqhllqltv wgikqlqari laverylkdq qllgiwgcsg
 541    arqllsgivq qqnnllraie aqqhllqltv wgikqlqari laverylkdq qllgiwgcsg

 601    klicttavpw naswsnksle qiwnnmtwme wdreinnyts lihslieesq nqqekneqel
 601    klicttavpw naswsnksle qiwnhttwme wdreinnyts lihslieesq nqqekneqel
 601    klicttavpw naswsnksle qiwnhttwme wdreinnyts lihslieesq nqqekneqel

 661    leldkwaslw nwfnitnwlw yiklfimivg glvglrivfa vlsvvnrvrq gysplsfqth
 661    leldkwaslw nwfnitnwlw yiklfimivg glvglrivfa vlsivnrvrq gysplsfqth
 661    leldkwaslw nwfnitnwlw yikifimivg glvglrivfa vlsivnrvrq ghsplsfqth

 721    lpiprgpdrp egieeegger drdrsirlvn gslaliwddl rslclfsyhr lrdlllivtr
 721    lptprgpdrp egieeegger drdrsirlvn gslaliwddl rslclfsyhr lrdlllivtr
 721    lptpggpdrp egieeegger drdrsirlvn gslaliwddl rslclfsyhr lrdlllivtr
```

```
781   ivellgrrgw ealkywwnll qywsqelkns avsllnatai avaegtdrvi evvqgayrai
781   ivellgrrgw ealkywwnll qywsqelkns avsllnatai avaegtdrvi evvqgacrai
781   ivellgrrgw ealkywwnll qywsqelkns avsllnatai avaegtdrvi evvqgacrai


841   rhiprrirqg lerill
841   rhiprrirqg lerill
841   rhiprrirqg lerill
```

JOURNAL OF VIROLOGY, Feb. 1987, p. 570–578
0022-538X/87/020570-09$02.00/0

Vol. 61, No. 2

Computer-Assisted Analysis of Envelope Protein Sequences of Seven Human Immunodeficiency Virus Isolates: Prediction of Antigenic Epitopes in Conserved and Variable Regions

SUSANNE MODROW,[1] BEATRICE H. HAHN,[2] GEORGE M. SHAW,[2] ROBERT C. GALLO,[3] FLOSSIE WONG-STAAL,[3] AND HANS WOLF[1]*

Max von Pettenkofer-Institut, 8000 Munich 2, Federal Republic of Germany[1]; University of Alabama, Birmingham, Alabama 25294[2]; and National Institutes of Health, Bethesda, Maryland 20892[3]

Received 16 June 1986/Accepted 15 October 1986

Independent isolates of human immunodeficiency virus (HIV) exhibit a striking genomic diversity, most of which is located in the viral envelope gene. Since this property of the HIV group of viruses may play an important role in the pathobiology of the virus, we analyzed the predicted amino acid sequences of the envelope proteins of seven different HIV strains, three of which represent sequential isolates from a single patient. By using a computer program that predicts the secondary protein structure and superimposes values for hydrophilicity, surface probability, and flexibility, we identified several potential antigenic epitopes in the envelope proteins of the seven different viruses. Interestingly, the majority of the predicted epitopes in the exterior envelope protein (gp120) were found in regions of high sequence variability which are interspersed with highly conserved regions among the independent viral isolates. A comparison of the sequential viral isolates revealed that changes concerning the secondary structure of the protein occurred only in regions which were predicted to be antigenic, predominantly in highly variable regions. The membrane-associated protein gp41 contains no highly variable regions; about 80% of the amino acids were found to be conserved, and only one hydrophilic area was identified as likely to be accessible to antibody recognition. These findings give insight into the secondary and possible tertiary structure of variant HIV envelope proteins and should facilitate experimental approaches directed toward the identification and fine mapping of HIV envelope proteins.

Most patients with acquired immunodeficiency syndrome (AIDS) or AIDS-related complex show specific antibodies directed against proteins of human immunodeficiency virus (HIV), which have virus-neutralizing activity (3, 37, 40, 46) and are supposed to be directed against antigenic determinants located on the surface glycoprotein, as has been shown for other enveloped virus particles. However, the virus seems to have adopted properties which allow evasion of the immune surveillance mechanisms of the host. Differences among various isolates of HIV have primarily been analyzed at the nucleotide sequence level (35) of independent isolates and also recently in sequential viral isolates from the same patient (21). These variations seem to be concentrated in the envelope protein-encoding region of HIV (20) and may be fundamentally important for the biology and pathogenicity of HIV (40, 50). For this reason and for the development of viral antigens for diagnostic or vaccine use, it is important to identify and characterize antigenic determinants located in the glycoprotein complex of HIV and to define their possible functions.

In this study, we analyzed the amino acid sequences of the envelope protein complexes derived from the nucleotide sequences of seven AIDS virus isolates (21, 36, 39, 40, 43), three of which represented sequential isolates from the same patient. The present work is an extension of previously published reports on the genetic variability of the HIV envelope protein complex, which mainly focused on the DNA and primary amino acid sequences (9, 28, 43). By computer analysis we predicted the secondary structure of gp120 and gp41, the cleavage products of gp160 (11), and predicted potential antigenic sites by superimposing this secondary structure with the values for hydrophilicity, flexibility, surface probability, and glycosylation. Thus, 11 potential antigenic sites were identified, 9 of which were located in the exterior part (gp120) of the envelope protein and 2 in the membrane-bound portion (gp41). Five highly variable regions were characterized, all contained in gp120, coinciding with the predicted epitopes. In sequential isolates from a single patient, all alterations of secondary structures occurred in those regions which were identified as antigenic epitopes.

These results indicate that genomic variations of the AIDS virus seem to be manifested mainly in the extracellular portion of the envelope protein. The fact that those variations coincide with possible antigenic sites suggests that these regions may be immunogenic and may be of fundamental importance for the pathobiology of the virus.

MATERIALS AND METHODS

DNA and protein sequences. The following convention will be used to designate the HIV strains used. Strains HTLV-III(BH10), LAV(1A), HTLV-III(HAT3), HTLV-III(WMJ1), HTLV-III(WMJ2), and HTLV-III(WMJ3) are referred to as BH10, LAV1A, HAT3, WMJ1, WMJ2, and WMJ3, respectively.

The nucleotide sequences of the envelope open reading frames of HIV strains BH10, LAV1A, and ARV2 have been previously reported (36, 39, 45). HAT3 is the nucleotide sequence of a virus isolated in 1983 from a patient with AIDS (20, 33, 43); WMJ1 is a virus isolate from a child with AIDS born in 1982 and infected perinatally by her HIV-positive mother. WMJ2 and WMJ3 are sequential isolates from this same patient taken 3 (WMJ2) and 7 (WMJ3) months after the first (21).

* Corresponding author.

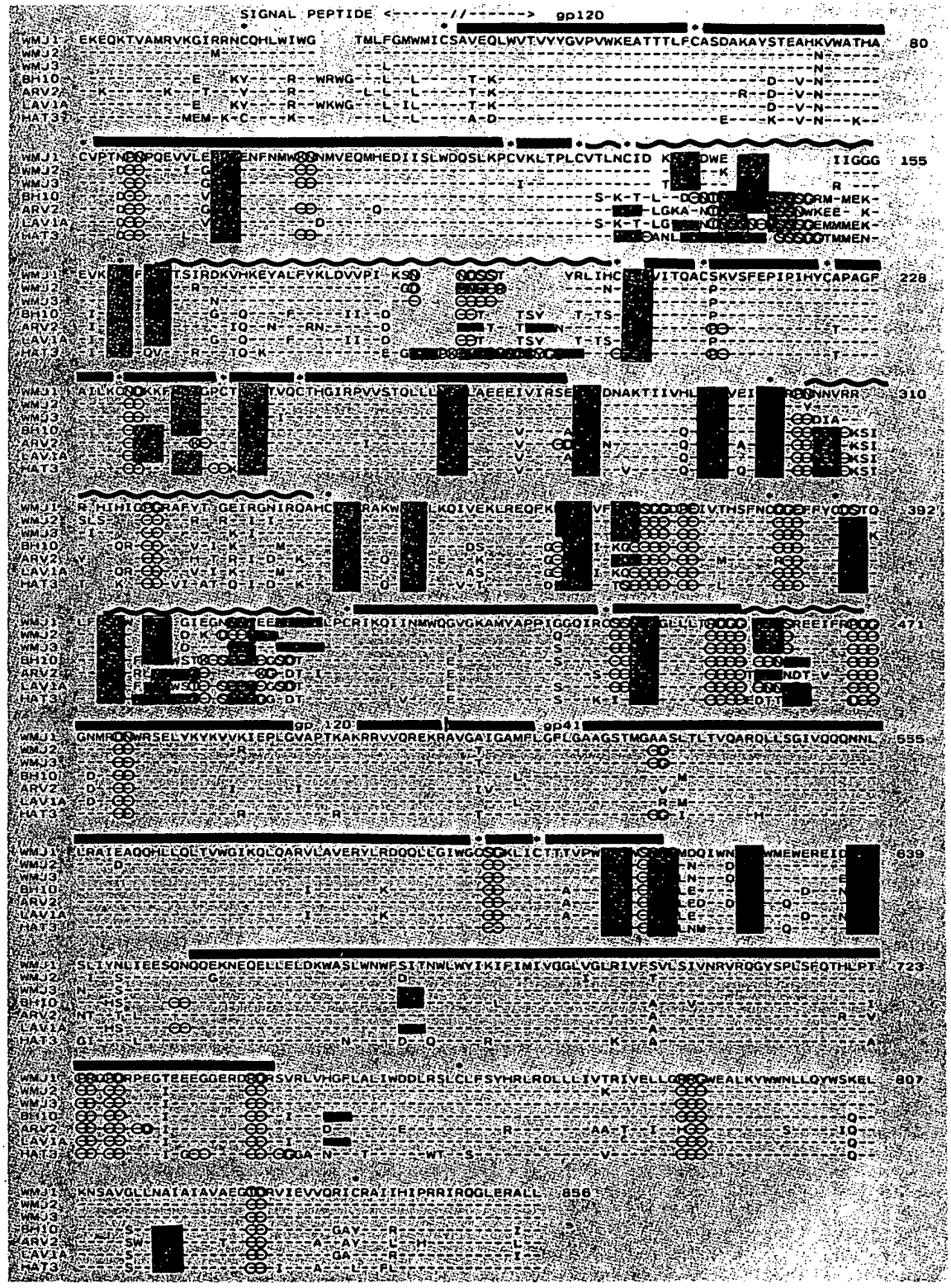


FIG. 1. Amino acid sequences of the entire envelope genes of seven HIV isolates. WMJ1, WMJ2, and WMJ3 are sequential isolates from the same patient; BH10, ARV2, LAV1A, and HAT3 represent independent virus isolates. Alignment of the sequences was done with the assistance of PRTLAN (49). Numbering of amino acid sequences is done from the first residue of the envelope protein open reading frame of WMJ1 up to the end. Dashes indicate amino acid identity with WMJ1, spaces indicate the absence of that amino acid. N-glycosylation sites are indicated as shaded areas, amino acids in β-turn regions are encircled. Symbols: *, cysteine residues; ▬, conserved regions; ~~~, variable regions.

Computer analysis of the sequences. The computer program of Queen and Korn (34) was used for translating nucleotide sequences into amino acids. Alignment of the deduced amino acid sequences of the surface glycoproteins of the seven viral isolates was accomplished with the computer program PRTALN (49). The secondary structure of the envelope proteins was predicted by a computer program written for a VAX750, based on suggestions by Cohen et al. (8) by using the algorithms of Chou and Fasman (6) or Garnier et al. (16). These predictions were superimposed with local hydrophilicity values (23). As an alternative to hydrophilicity, the values for surface probability (modified from Emini et al. [13]) or flexibility (24) were superimposed. For Chou-Fasman calculations, the probability of the occurrence of α-helixes, β-pleated sheets, β-turn regions, and random coil structures were evaluated by using stringent conditions: $P_{boundary} \geq 1$, with $P_\beta > P$ and $P_t > P_\alpha$ (P_β = probability for β-sheet, P_α = probability for α-helical region, and P_t = probability for β-turn regions).

The parameters were averaged over five amino acid residues with a limit of 0.7 for hydrophilicity, 5.0 for surface probability, and 1.040 for flexibility. β-Turn regions adjacent to β-sheets or α-helical regions in a hydrophilic or nonhydrophobic environment combined with the simultaneous occurrence of high values for flexibility and surface probability in those regions were considered to be candidates for potential antigenicity. These structures are believed to form freely accessible and mobile loops at the protein surface and are thus considered to be prime candidates for antigenic sites (2, 6, 12, 23, 28, 47).

RESULTS

Comparison of amino acid sequences. The open reading frames of the envelope gene derived from the nucleotide sequence of the seven virus isolates are similar in size and encode 854 (WMJ3) to 873 (HAT3) amino acids (Fig. 1). Methionine codons at positions 8 (BH10 and LAV1A) or 9 (HAT3, ARV2, WMJ1, WMJ2, and WMJ3) supposedly mark the beginning of the envelope protein with a potential leader sequence, which is cleaved from the envelope precursor protein during maturation (1). The peptide sequences from position 38 represent the envelope precursor gp160, which is cleaved into the exterior gp120 (with 23 to 25 potential N-glycosylation sites) and the membrane-bound gp41 (containing 4 to 7 potential N-glycosylation sites). A stretch of positively charged amino acids at positions 510, 509, 508 (WMJ1, WMJ2, WMJ3), 517 (ARV2), 518 (BH10), 523 (LAV1A), and 527 (HAT3), respectively, marks the cleavage site (1, 11).

Although the overall sizes and structures of the seven surface proteins are rather similar, the deduced amino acid sequences differ substantially. On the average only 66% of the amino acids are conserved in the exterior part of the protein, and these changes are clustered in special regions with only up to 10% conserved amino acids. gp41, the transmembrane part of the envelope protein complex, shows more than 80% conserved amino acids (Fig. 1, Table 1) and no regions of high variability. Furthermore, the changes in the latter region are all due to point mutations, whereas changes in gp120 frequently result from insertions and deletions and appear as clustered mutations interspersed with segments which have a high content of conserved amino acids (89%).

According to their content of conserved amino acids, glycosylation sites, and β-turn regions, the envelope proteins of HIV were subdivided into highly variable and constant regions. Constant regions were defined to contain 75% or more conserved amino acids and to have no amino acid insertions and deletions. In contrast, variable regions showed a low degree of conserved residues (25% or less) and a great variability in length due to deletions or insertions; these changes occurred at at least every fifth amino acid.

Due to the high number of conserved amino acid residues in the constant regions, glycosylation sites and secondary structures also showed a low degree of variation; more than 50% of potential glycosylation sites and amino acids in β-turn configurations were found to be conserved. Regions of high variability had a low degree of conserved β-turns (0 to 25%), especially in a hydrophilic environment, and almost no conserved potential glycosylation sites (0 to 25%).

By these criteria, surface glycoprotein gp160 could be subdivided into a clustered pattern of highly variable (V1 to V5) and constant regions (C1 to C6). For purposes of classification the minimum length of a region was set at 10 amino acid residues; the locations and parameters of those regions are shown in Fig. 1 and 2 and Table 1.

Prediction of antigenic determinants. (i) Exterior envelope protein gp120. For further analysis we predicted antigenic determinants in the amino acid sequences with a computer program which predicted the secondary structure and calculated the values for hydrophilicity, flexibility, and surface probability (Fig. 2, Table 1). These regions are mainly located in β-turn regions which show a high degree of nonhydrophobic or flexible amino acid sequences or are predicted to have a high probability of location at the surface of the polypeptide (8, 12, 13, 19, 32, 44).

In gp120, nine epitopes (I to IX) with a high antigenic potential can be predicted (Table 2, Fig. 2). Epitope I is located in a region of high variability (V1) at amino acids 137 to 154 of viral strain WMJ1. The locations of the antigenic epitopes for the other isolates are indicated in Table 2.

Epitope I has elevated values for hydrophilicity and flexibility, 1 to 3 potential glycosylation sites, and a number (up to 11 in LAV1A) of amino acids in β-turn configurations. Isolates WMJ1 and WMJ3 show no β-turns due to the high variability of this region. Epitope II (amino acids 186 to 203) is situated in variable region V2 and contains residues with high values for hydrophilicity, flexibility, and surface probability; up to 15 amino acids (HAT3) have β-turn configurations (with the exception of ARV2), and all strains may be glycosylated. Epitope III (amino acids 232 to 246) in constant region C2 shows two conserved hydrophilic β-turns, two conserved potential glycosylation sites, and high values of flexibility. Epitope IV (amino acids 300 to 320) in variable region V3 shows 2 to 8 amino acids in β-turns, with elevated values for hydrophilicity, flexibility, and surface probability, and has 1 to 2 potential glycosylation sites. Epitope V (amino acids 358 to 375) is only slightly variable (72% conserved amino acids), has high values for hydrophilicity, surface probability, and flexibility, and has 1 to 2 N-glycosylation sites and 6 to 8 residues in β-turn configurations. Epitope VI (amino acids 394 to 412), in contrast, is in highly variable region V4, with 2 to 4 possible glycosylation sites, 4 to 11 β-turns, and high values for hydrophilicity and flexibility; the values for surface probability are not elevated. Epitope VII (amino acids 445 to 458) in constant region C3 has 4 to 7 β-turns. One glycosylation site is hydrophilic and flexible and is directly followed by epitope VIII (amino acids 459 to 469), which shows high variability (V5), is hydrophilic, flexible, and possibly glycosylated, and has one (ARV2) to six residues in β-turn configurations. Epitope IX (amino

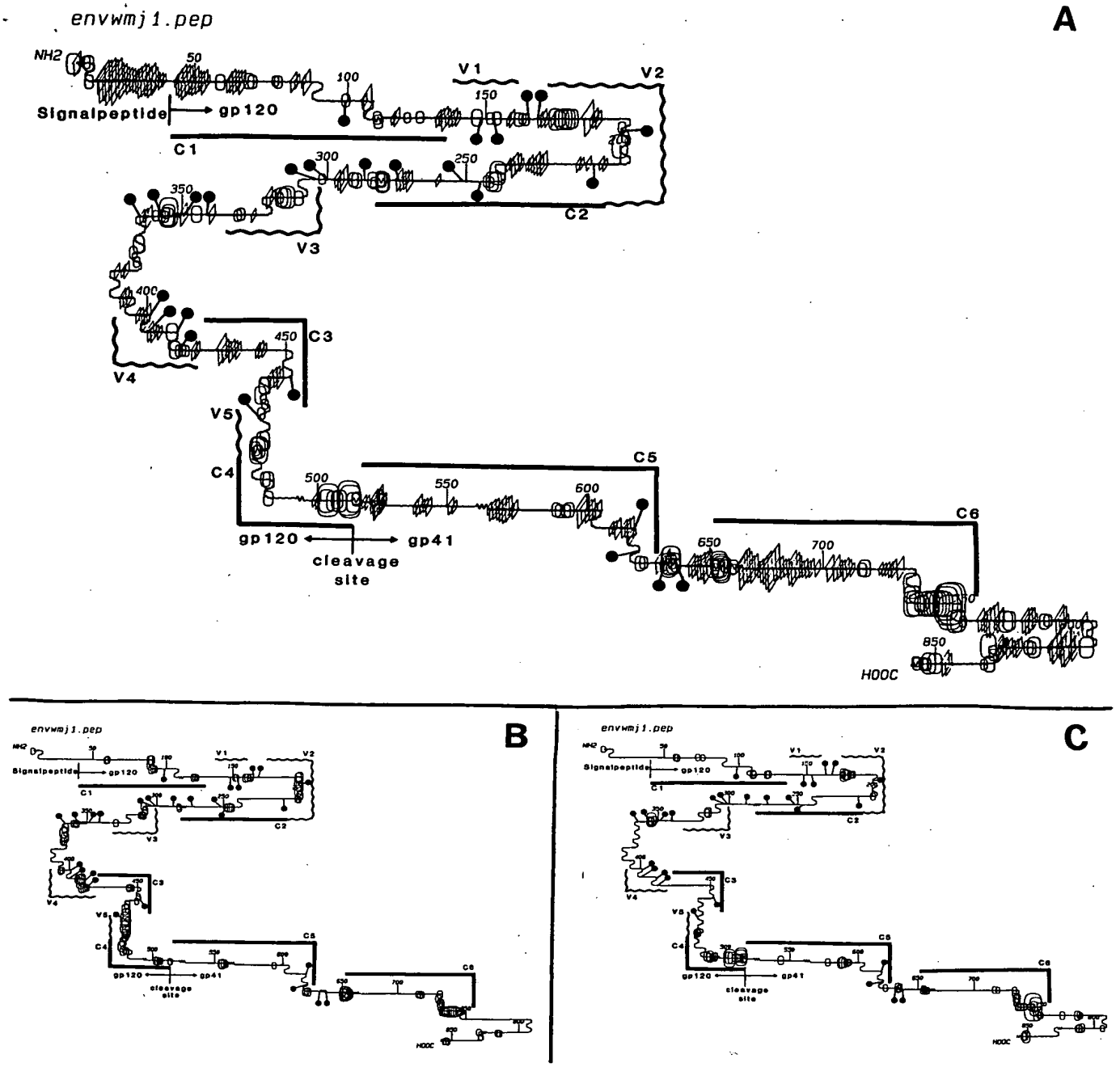


FIG. 2. Chou-Fasman prediction of the envelope protein gp160 derived from the sequence of WMJ1. The computer plots start at the methionine at residue 9 of the open reading frame. The probability of the occurrence of α-helices (~~~), β-pleated sheets (~~), random coils (~~), and β-turn regions (ↄ) were evaluated by using stringent conditions. The parameters for hydrophilicity, flexibility, and surface probability were averaged over five amino acid residues, with a limit of 0.7 for hydrophilicity, 1.040 for flexibility, and 5.0 for surface probability. Symbols: ●—, N-glycosylation sites; ▬, conserved regions; ~~~, variable regions. (A) Secondary structures superimposed with the values for hydrophilicity; 0, hydrophilic regions, ◇, hydrophobic regions. (B) Secondary structures superimposed with the values for flexibility; 0, flexible regions. (C) Secondary structures superimposed with the values for surface probability; 0, high surface probability.

acids 470 to 483) in region C4 has high values for all parameters, four conserved β-turns, and no potential glycosylation sites. In fact, epitopes VII to IX are one continuous antigenic region which is subdivided because constant regions C3 and C4 are interrupted by V5, a rather short region with high variability.

There may be additional epitopes in singular virus isolates (e.g., amino acids 285 to 293 in ARV2) which were not classified because similar epitopes could not be found in the other virus strains.

Some of these epitopes must be considered of lesser antigenic quality, since they have only a few β-turns (epitope

TABLE 1. Parameters and values for conserved and variable regions, calculated from the sequences of seven HIV envelope proteins (gp 160)[a]

Region	Length	Amino acid deletions/ insertions	Conserved amino acids		Glycosylation sites	Conserved glycosylation sites		β-Turns	Conserved β-turns	
			No.	%		No.	%		No.	%
C1	97/97/97	0	84	87	1/1/1	1	100	4/4/4	2	50
38–134	97/97/97/97				1/1/1/1			2/4/2/4		
V1	20/20/20	0–15	3	10	2/2/2	0	0	0/2/0	0	0
135–154	24/23/29/25				1/2/3/3			8/5/11/7		
	8/8/8	0	7	87	1/1/1	1	100	0/0/0	0	100
155–162	8/8/8/8				1/1/1/1			0/0/0/0		
V2	41/41/41	0–18	11	21	2/3/2	1	20	5/6/5	0	0
163–203	39/44/39/52				2/4/3/5			2/2/2/17		
C2	76/76/76	0	68	89	3/3/3	2	50	2/2/2	2	50
204–279	76/76/76/76				4/3/4/3			2/4/2/4		
	25/25/25	0	19	76	3/3/3	3	100	0/0/0	0	0
280–304	25/25/25/25				3/3/3/3			0/3/0/0		
V3	26/26/26	0–5	7	25	0/0/0	0	0	4/2/4	2	25
305–330	27/26/27/26				1/1/1/1			8/6/8/8		
	65/65/65	0–1	48	72	5/6/6	4	67	11/9/10	8	72
331–395	66/66/66/65				5/5/5/6			11/10/11/10		
V4	19/19/19	0–11	4	17	3/2/3	0	0	4/6/4	0	0
396–414	23/18/23/19				2/1/2/3			11/4/11/9		
C3	44/44/44	0	38	86	1/1/1	1	100	7/7/7	4	57
415–458	44/44/44/44				1/1/1/1			7/7/7/4		
V5	11/11/11	0–2	2	16	1/1/1	0	0	5/5/5	1	18
459–469	11/12/11/12				1/1/1/1			5/1/6/1		
C4	41/41/41	0	36	87	0/0/0	0	100	4/4/4	4	100
470–510	41/41/41/41				0/0/0/0			4/4/4/4		
C5	106/106/106	0–1	95	90	2/2/2	2	100	6/8/8	6	75
511–616	105/106/105/106				2/2/2/2			6/6/6/8		
	38/38/38	0	17	44	2/2/2	2	100	2/0/0	0	0
616–653	38/38/38/38				2/2/2/2			2/0/2/2		
C6	92/92/92	0	84	85	0/0/1	0	0	6/6/6	6	66
654–745	92/92/92/92				1/0/1/0			6/8/6/10		
	111/111/111	0	77	69	0/0/0	0	0	5/5/5	4	57
746–856	111/111/11/111				2/1/2/2			5/4/5/7		

III) or no high values for surface probability (epitopes I, VI, VII, and VIII).

(ii) Transmembrane protein gp41. Only two antigenic sites could be identified in gp41. These contain amino acids 612 to 635 and 722 to 745 (Fig. 1 and 2, Tables 1 and 2). The first, epitope X, is located in a slightly variable region (56% conserved amino acids) and contains four conserved glycosylation sites, has high values for hydrophilicity, surface probability, and flexibility, has 4 to 6 β-turns, and is probably the only antigenic site in gp41 which is located outside the lipid bilayer and accessible to antibody reaction and recognition. The second epitope (amino acids 722 to 745) is located directly after a stretch of hydrophobic amino acids which is likely to be a transmembrane region of gp41 (TM3); these hydrophilic, flexible amino acids in β-turn regions with high surface probability probably represent the hydrophilic anchor sequence which has been identified in most membrane proteins (10). This region, however, should be inside the cell and thus is not the best epitope for antigenic response. The following stretch of about 100 amino acids has a further region which might be an antigenic determinant. This region, whose function is not known, might however be gradually cleaved off by proteases from the precursor before maturation of gp41. With antibodies against a synthetic peptide derived from this region (25), mainly the precursor gp160 could be identified; furthermore, the molecular size of a glycosylated gp41 should be about 52,000 to 54,000 daltons by calculation (42,000 daltons of the primary product plus the molecular size of the carbohydrate moiety) unless it is proteolytically processed.

Alterations in the envelope proteins of the sequential isolates. The envelope proteins of the sequential viral isolates from the same patient which were taken at intervals of 3 and 4 months are very similar in size and show only three amino acid deletions or insertions. Most changes are due to point mutations (21). Most of the alterations in the amino acid sequences which are due to those mutations are located in those regions which were found to be highly variable. Differences in the secondary structures (Fig. 1) were identified in regions V1, V2, V3, and V4, corresponding with antigenic epitopes I, II, IV, and VI, and in epitopes V and X, which are located in slightly variable regions (76 and 56% of conserved amino acids, respectively). A further β-turn alteration occurs in region C5 in a short hydrophilic environment which may be located between two transmembrane-spanning regions (Fig. 1, 2, and 3). That means that all alterations concerning the secondary structure of the envelope protein, with the exception of that in region C5, are located in the predicted antigenic determinants; epitopes III, VII, VIII, IX, and XI are conserved. Interestingly, some of the β-turn alterations identified in isolates WMJ1 and WMJ2 are reversed in isolate WMJ3 (I, II, IV, and VI), even if there are further variations in the amino acid sequence (I and IV).

Prediction of tertiary structures. After cleavage of the leader peptide from precursor protein gp160, gp120 represents the highly glycosylated exterior part of the envelope

TABLE 1—*(Continued)*

Region	Length	Hydrophilic β-turns	Conserved hydrophilic β-turns		Hydrophilicity		Surface probability		Flexibility	
			No.	%	No.	%	No.	%	No.	%
C1	97/97/97	2/2/2	2	100	10/8/7	7–10	7/7/9	7–10	11/1/11	8–10
38–134	97/97/97/97	2/2/2/2			8/9/9/9		8/9/7/10		11/11/8/12	
V1	20/20/20	0/0/0	0	0	5/5/7	13–33	0/0/0	0–13	6/7/6	28–82
135–154	24/23/29/25	3/0/1/0			9/9/4/4		2/3/3/0		14/19/17/14	
	8/8/8	0/0/0	0	100	0/0/0	0	0/0/0	0	0/0/0	0
	8/8/8/8	0/0/0/0			0/0/0/0		0/0/0/0		0/0/0/0	
V2	41/41/41	4/4/4	0	0	9/17/6	14–41	9/12/4	10–32	8/11/12	20–46
163–203	39/44/39/52	2/0/2/3			7/9/7/9		5/5/5/12		11/9/11/24	
C2	76/76/76	2/2/2	2	67	9/8/9	7–11	0/0/0	0	5/5/5	7–11
204–279	76/76/76/76	2/3/2/2			5/5/5/9		0/0/0/0		7/8/9/6	
	25/25/25	0/0/0	0	0	5/5/5	8–20	0/0/0	0	2/2/2	8–12
280–304	25/25/25/25	0/2/0/0			2/2/2/4		0/0/0/0		0/3/0/2	
V3	26/26/26	3/1/3	1	16	7/5/8	19–34	2/5/5	7–22	5/0/3	0–57
305–330	27/36/27/26	6/7/6/4			7/7/7/9		6/2/6/3		12/8/12/15	
	65/65/65	4/4/4	4	57	11/11/11	12–17	5/5/5	6–8	8/8/9	12–25
331–395	66/66/66/65	6/7/6/4			11/8/9/10		5/5/4/4		17/10/10/10	
V4	19/19/19	1/4/1	0	0	5/7/7	2–11	0/0/0	0–6	9/13/9	10–20
396–414	23/18/23/19	8/2/6/3			3/4/3/1		0/1/0/0		12/7/12/13	
C3	44/44/44	4/4/4	4	100	0/1/0	0–11	0/0/0	0	5/5/5	5–11
415–458	44/44/44/44	4/4/4/4			2/3/5/2		0/0/0/0		5/2/5/4	
V5	11/11/11	2/2/2	0	0	10/10/10	1–90	4/5/5	0–45	10/6/10	54–100
459–469	11/12/11/12	1/0/1/0			8/2/1/7		0/0/0/1		11/9/9/11	
C4	41/41/41	4/4/4	4	100	12/12/12	29–32	18/19/19	43–46	12/11/12	24–29
470–510	41/41/41/41	4/4/4/4			13/13/13/12		19/19/19/19		12/12/12/10	
C5	106/106/106	0/0/0	0	100	5/5/5	5	8/8/8	5–7	5/5/5	6
511–616	105/106/105/106	0/0/0/0			5/5/5/5		8/8/8/5		6/5/6/5	
	38/38/38	0/0/0	0	0	7/5/5	13–23	4/4/4	10–13	3/3/3	7–10
616–653	38/38/38/38	2/0/2/2			7/5/7/4		3/4/3/4		3/3/3/2	
C6	92/92/92	6/6/6	6	67	34/34/34	34	29/20/23	20–29	35/35/36	32–40
654–745	92/92/92/92	6/8/6/10			34/34/34/34		19/22/24/28		33/32/37/40	
	111/111/111	5/5/5	2	39	20/20/19	15–18	13/13/13	10–12	9/9/9	7–10
746–856	111/111/111/111	5/2/5/7			20/17/20/20		12/12/12/11		10/8/10/11	

[a]Values are arranged in the following mode: WMJ1/ WMJ2/ WMJ3
BH10/ ARV2/ LAV1A/HAT3.

protein complex of HIV. This polypeptide part is likely connected only via the ionic interactions of 14 positively charged amino acids in the carboxy-terminal region of gp120 (region C4) adjacent to the cleavage site with the negatively charged phosphate groups in the bilayer membrane. There are 18 conserved cysteine residues dispersed over the gp120 sequence (Fig. 1) which might also be involved in the complex formation of gp120 and gp41. A report describing the frequent loss of gp120 from the particles during purification and immunoelectron microscopy argues, however, against a direct involvement of disulfite bonds in the complex formation of gp120 and gp41 (17).

The transmembrane polypeptide gp41 contains three stretches of hydrophobic amino acids. Directly after the cleavage site there are about 60 hydrophobic residues (amino acids 511 to 571) which are interrupted by a stretch of about 10 amino acids that are predicted to have a high surface probability and, in comparison with the surrounding residues, a higher hydrophilicity and flexibility and some alterations in β-turns. From the patterns of other viral (22, 32) or

TABLE 2. Antigenic epitopes in the various strains

Epitope no.	Region	Amino acid residues in virus strain:						
		WMJ1	WMJ2	WMJ3	BH10	ARV2	LAV1A	HAT3
I	V1	137–154	137–154	137–153	137–157	137–158	137–163	137–158
II	V2	186–203	186–203	185–202	188–203	189–209	195–209	189–217
III	C2	232–246	232–246	228–245	232–246	240–253	240–253	249–261
IV	V3	300–320	300–319	299–318	300–321	300–327	308–328	316–355
V		358–375	257–374	356–373	360–377	367–384	367–384	375–391
VI	V4	394–412	393–411	392–410	397–418	404–420	404–425	410–427
VII	C3	445–458	444–457	443–456	451–464	453–466	457–470	459–472
VIII	V5	459–469	458–468	457–467	465–475	467–478	471–481	473–484
IX	C4	470–483	469–482	468–481	476–489	478–492	482–496	485–499
X		611–637	610–636	609–635	616–643	620–646	623–649	627–653
XI	C6	724–745	723–744	722–743	729–750	733–754	736–757	740–761

membrane proteins, e.g., bacteriorhodopsin (14), acetylcholine receptor (15), and erythrocyte band III (26), such regions are known to represent sequential transmembrane-spanning regions (TM1 and TM2 [Fig. 3]). A domain of hydrophilic, flexible, and charged amino acids (amino acids to 572 to 594) linked to the hydrophobic region may represent a hydrophilic anchor sequence.

Transmembrane region TM3, which was suggested from amino acids 670 to 695, with its hydrophilic anchor has already been mentioned. In transmembrane regions TM2 and TM3 three charged amino acids (arginine at positions 557 and 707 and glutamic acid at position 560) are identifiable. Those charged amino acids were also reported in other transmembrane regions in multi-spanning proteins and do not lead to disturbance of the overall hydrophobic character of a transmembrane region (48). Between TM2 and TM3 hydrophobic regions from the N terminus of gp41 are 90 amino acid residues (580 to 670) which should be facing the outer side of the lipid bilayer. Within this region are four conserved glycosylation sites and a good antigenic probability (epitope X). These considerations were combined into a three-dimensional model of the HIV envelope protein complex (Fig. 3). Alternatively, TM1 and TM2 may represent an apolar hydrophobic stretch of amino acids located outside the viral membrane and may play a role similar to the paramyxovirus fusion protein by penetrating into the membrane of the fusion partner (31). The same could occur if the C-terminal part of gp41, including or not including TM3, would be directed to the outer membrane or be in equilibrium between those two orientations. This possibility is supported by the observation that antiserum against a peptide from amino acids 725 to 752 gives cell surface labeling and is neutralizing (5a). The predicted antigenic determinants are not affected by the discussed models.

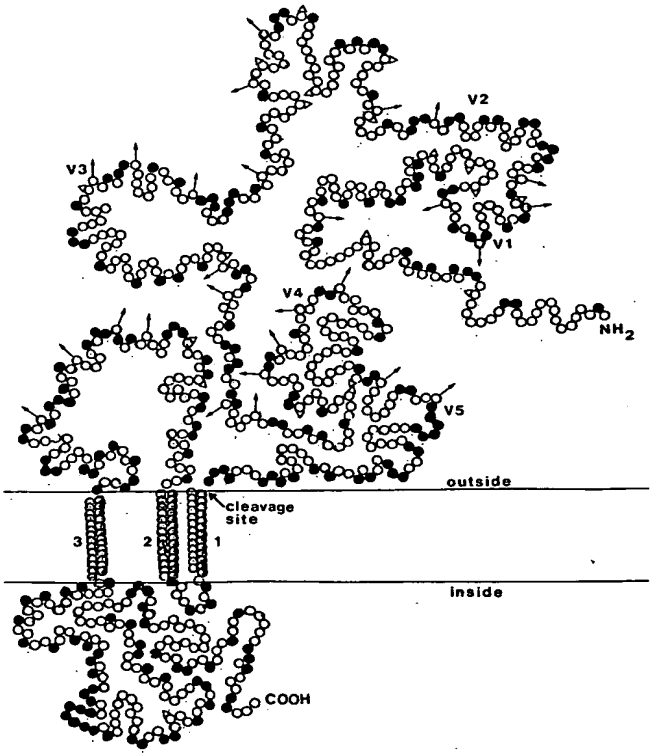


FIG. 3. Suggestion for the tertiary structure of envelope protein gp160 beginning at amino acid 38. Each circle represents one amino acid residue, shaded areas indicate transmembrane-spanning regions. Variable regions are indicated. ●, Polar amino acids; arrows, N-glycosylation sites; ⬡, cysteine residues.

DISCUSSION

Previous studies have shown that variability in the genomes of different HIV isolates is a prominent feature of this group of viruses (21, 40, 43, 50). In the present study we analyzed the deduced amino acid sequences of seven HIV isolates and predicted the secondary structures in combination with the values for hydrophilicity, flexibility, and surface probability, which resulted in the identification of potentially antigenic epitopes.

Analysis of the primary and secondary structure of polypeptide shows that antigenic determinants are often found in loop like structures on the surface of a molecule that have been identified as biologically important regions in several other viral membrane or capsid proteins (8, 12, 19, 28, 30, 32, 47). Antigenic sites formed by the folding of the amino acid chain to tertiary structures contribute to immunological activities; this cannot be predicted by calculations similar to those applied here. By our methods, 11 epitopes could be identified and characterized in the envelope protein complex of HIV, 9 of which are located in the exterior protein gp120. A comparison of the amino acid sequences of the seven viral strains led to the subdivision of gp120 into highly variable and conserved regions. Five highly variable regions could be identified, all of which coincided with the predicted epitopes I, II, IV, VI, and VIII. These epitopes, due to their concentration of β-turns and hydrophilic amino acids, have a very high potential for antigenicity. Genomic variation primarily occurs in the exterior envelope sequences and corresponds to predicted epitopes of different AIDS virus strains, which differ also in sequential isolates from the same patient (epitopes I, II, IV, and VI), suggesting that immune selection may play an important role in the generation of variant virus strains. Two related retroviruses, equine infectious anemia virus and visna virus (18, 41, 42), show similar progressive changes in their envelope genes. There is evidence that these changes do, in fact, lead to substantial changes in the antigenic properties of the envelope which may reflect immune selective pressure exerted by the host (7, 29, 38).

Whereas the variable regions of the exterior protein gp120 possess properties typical for antigenic sites, the conserved regions are generally hydrophobic and lack areas with a high number of β-turns. There are only three exceptions: epitope III in region C2, epitope VII in region C3, and epitope IX in region C4. Epitope III has only two amino acids predicted in β-turn configurations and must be considered of minor antigenic potential, in comparison with the other epitopes. Epitope VII shows many β-turns; these however are only slightly hydrophilic and have no high values for surface probability. Epitope XI in C4 (amino acids 470 to 483) consists of a stretch of hydrophilic, flexible amino acids in β-turn areas with high surface probability; this region is adjacent to the cleavage site, and a synthetic peptide corresponding to that region reacted with about 80% of the sera from HIV-positive individuals (S. Modrow, unpublished data). However, it is unclear if this epitope is conserved in all virus strains. It is also possible that conserved region C4 might contribute to biologically important functions of the virus particle such as virus cell adsorption and connection of

gp120 to the envelope membrane; both functions could protect this region from selective immune pressure.

The highly conserved protein structures together with 18 cysteine residues conserved in the seven strains may contribute to the stability of a constant core structure of gp120, with variable, highly antigenic regions looping out.

In contrast to the exterior protein gp120, transmembrane polypeptide gp41 contains no region of high variability and consists mainly of regions with hydrophobic areas which may be arranged into three transmembrane regions. Between these transmembrane regions is a glycosylated region with good predicted antigenicity (epitope X), which may be located outside the viral envelope and thus may be recognized by the immune system of the host. This region has only slight variability and may contribute to biologically important functions as well. It thus could represent a valuable diagnostic antigen. It has been shown (4, 5) that segments of the protein, including epitope X, when produced as recombinant gene products in bacteria, are recognized by HIV-positive sera in enzyme-linked immunosorbent assays and Western blots; preliminary results with a synthetic peptide from this region give similar results (Modrow, unpublished). The last epitope, epitope XI, consists of stretches of highly hydrophilic, flexible amino acids with high values for surface probability and several β-turns. This region may represent the hydrophilic anchor sequences located adjacent to transmembrane region TM3 and thus should be located inside the cell.

The observation that HIV binds to the CD4 molecule of T cells (27) which is supposed to bind to molecules of the major histocompatibility complex class II during stimulation of immune responses may suggest functional similarities of yet unidentified regions of gp120 or gp41 with those of the major histocompatibility complex class II molecules. Our analysis of constant and variable regions may well identify candidates for structures of dominant immunogenicity for humoral immune response. This does not, however, exclude the possibility that during biological degradation of viral structures other segments, including constant regions of the viral envelope proteins, lead to formation of specific antibody responses in patients. Such antibodies, when detected by in vitro antigens not presented in their native configuration, can be of diagnostic value. In addition, these antibodies may also have neutralizing activity, since the respective sequences recognized may be located in accessible areas of the envelope or be accessible during molecular mobility.

Cellular immune mechanisms have not been addressed so far and are not predictable by our protein analysis. However, in this case, sequential rather than complex antigens seem to be of particular importance. Selective priming of cell-mediated immunity with epitopes selected from detailed analysis of the viral envelope should open new approaches for the control of HIV-related disease.

LITERATURE CITED

1. **Allen J. S., J. E. Coligan, F. Barin, M. F. McLane, J. G. Sodroski, C. A. Rosen, W. A. Haseltine, T. H. Lee, and M. Essex.** 1985. Major glycoprotein antigens that induce antibodies in AIDS patients are encoded by HTLV-III. Science 228:1091–1093.
2. **Atassi, M. Z.** 1978. Precise prediction of the entire antigenic structure of lysozyme. Molecular features of protein antigenic structures and potential of "surface stimulation" synthesis—a powerful new concept for protein binding sites. Immunochemistry 15:909–936.
3. **Barin, F., M. F. McLane, J. S. Allan, T. H. Lee, J. E. Groopman, and M. Essex.** 1985. Virus envelope protein of HTLV-III represents major target antigen for antibodies in AIDS-patients. Science 228:1094–1096.
4. **Cabradilla, C. D., J. E. Groopman, J. Lanigan, M. Renz, L. A. Laskey, and D. J. Capon.** 1986. Serodiagnosis of antibodies to the human AIDS retrovirus with a bacterially synthesized env polypeptide. Biotechnology 4:128–133.
5. **Chang, T. W., K. Ikunoshin, S. McKinney, P. Chanda, A. D. Barone, F. Wong-Staal, R. C. Gallo, and N. T. Chang.** 1985. Detection of antibodies to human T-cell lymphotropic virus III (HTLV-III) with an immunoassay employing a recombinant *Escherichia coli*-derived viral antigenic peptide. Biotechnology 3:905–909.

5a. **Chanh, T. C., G. R. Dreesman, P. Kanda, G. Linette, J. T. Sparrow, D. D. Ho, and R. C. Kennedy.** 1986. Induction of anti-HIV neutralizing antibodies by synthetic peptides. EMBO J. 5:3065–3073.

6. **Chou, P. Y., and G. D. Fasman.** 1974. Prediction of protein conformation. Biochemistry 13:222–245.
7. **Clements, Y. E., F. S. Pedersen, O. Narayan, and W. A. Haseltine.** 1980. Genomic changes associated with antigenic variation of visna virus during persistent infection. Proc. Natl. Acad. Sci. USA 77:4454–4458.
8. **Cohen, G. H., B. Dietzschold, M. Ponce de Leon, D. Long, E. Golub, A. Varrichio, L. Pereira, and R. J. Eisenberg.** 1984. Localization and synthesis of an antigenic determinant of herpes simplex virus glycoprotein D that stimulates the production of neutralizing antibody. J. Virol. 49:102–108.
9. **Crowl, R., K. Ganguly, M. Gordon, R. Conroy, M. Schaber, R. Kramer, G. Shaw, F. Wong-Staal, and E. P. Reddy.** 1985. HTLV-III *env* gene products synthesized in *E. coli* are recognized by antibodies present in the sera of AIDS-patients. Cell 41:979–986.
10. **Davis, N. G., J. D. Boeke, and P. Model.** 1985. Fine structure of a membrane anchor region domain. J. Mol. Biol. 181:111–121.
11. **DiMarzo-Veronese, F., A. L. deVico, T. D. Copeland, S. Oroszlan, R. C. Gallo, and M. G. Sarngadharan.** 1985. Characterization of gp41 as the transmembrane protein coded by the HTLV-III/LAV envelope gene. Science 229:1402–1405.
12. **Eisenberg, R. J., D. Long, M. Ponce de Leon, J. T. Matthews, P. G. Spear, M. G. Gibson, L. A. Lasky, P. Berman, E. Golub, and G. H. Cohen.** 1985. Localization of epitopes of herpes simplex virus type 1 glycoprotein D. J. Virol. 53:634–644.
13. **Emini, E. A., J. V. Hughes, D. S. Perlow, and J. Boger.** 1985. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J. Virol. 55:836–839.
14. **Engleman, D., A. Goldman, and T. Steitz.** 1982. The identification of helical segments in the polypeptide chain of bacteriorhodopsin. Methods Enzymol. 88:81–98.
15. **Finer-Moore, J., and R. M. Stroud.** 1984. Amphipathic analysis and possible formation of the ion channel in acetylcholine receptor. Proc. Natl. Acad. Sci. USA 81:155–159.
16. **Garnier, J., D. J. Osguthorpe, and B. Robson.** 1978. Analysis of the accuracy and implication of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. 120:97–120.
17. **Gelderblom, H. R., H. Reupke, and G. Pauli.** 1985. Loss of envelope antigens of HTLV-III/LAV, a factor in AIDS-pathogenesis? Lancet ii:1016–1017.
18. **Gonda, M. A., F. Wong-Staal, R. C. Gallo, J. E. Clements, O. Narayan, and R. V. Gilden.** 1985. Sequence homology and morphologic similarities of HTLV-III and visna virus, a pathogenic lentivirus. Science 227:173–177.
19. **Gunn, P. R., F. Sato, K. F. H. Powell, A. R. Bellamy, J. R. Napier, D. R. K. Harding, W. S. Hancock, L. J. Siegman, and G. W. Both.** 1985. Rotavirus neutralizing protein VP7: antigenic determinants investigated by sequence analysis and peptide synthesis. J. Virol. 54:791–797.
20. **Hahn, B. H., M. A. Gonda, G. M. Shaw, M. Popovic, J. Hoxie, R. C. Gallo, and F. Wong-Staal.** 1985. Genomic diversity of the AIDS virus HTLV-III: different viruses exhibit greatest divergence in their envelope genes. Proc. Natl. Acad. Sci. USA 82:4813–4817.
21. **Hahn, B. H., G. M. Shaw, M. E. Taylor, R. R. Redfield, P. D.**

Markham, S. Z. Salahuddin, F. Wong-Staal, R. C. Gallo, E. S. Parks, and W. P. Parks. 1986. Genetic variation in HTLV-III/LAV over time in patients with AIDS or at risk of AIDS. Science 232:1548–1554.
22. Hennessey, K., S. Fennewald, M. Hummel, T. Cole, and E. Kieff. 1984. A membrane protein encoded by Epstein-Barr virus in latent growth-transforming infection. Proc. Natl. Acad. Sci. USA 81:7207–7211.
23. Hopp, T. P., and K. R. Woods. 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA 78:3824–3828.
24. Karplus, P. A., and G. E. Schulz. 1985. Prediction of chain flexibility in proteins. Naturwissenschaften 72:212–213.
25. Kennedy, R. C., R. D. Henkel, D. Pauletti, Y. S. Allan, T. H. Lee, M. Essex, and G. R. Dreesman. 1986. Antiserum to a synthetic peptide recognizes the HTLV-III envelope glycoprotein. Science 231:1556–1559.
26. Kopito, R., and H. F. Lodish. 1985. Primary structure and transmembrane orientation of the murine anion and exchange protein. Nature (London) 316:234–238.
27. McDougal, J. S., M. S. Kennedy, J. M. Sligh, S. P. Cort, A. Mawle, and J. K. A. Nicholson. 1986. Binding of HTLV-III/LAV to T4$^+$ T-cells by a complex of the 110kD viral protein and the T4 molecule. Science 231:382–385.
28. Modrow, S., and H. Wolf. 1986. Characterization of two related Epstein-Barr virus-encoded membrane proteins that are differentially expressed in Burkitt lymphoma and in vitro transformed cell lines. Proc. Natl. Acad. Sci. USA 83:5703–5707.
29. Montelaro, R. C., B. Parekh, A. Orego, and C. J. Issel. 1984. Antigenic variation during persistent infection by equine infectious anemia virus and retrovirus. J. Biol. Chem. 259: 10539–10544.
30. Motz, M., J. Fan, R. Seibl, W. Jilg, and H. Wolf. 1986. Expression of the Epstein-Barr virus 138 kDa early protein in *Escherichia coli* for the use as antigen in diagnostic tests. Gene 42:303–312.
31. Paterson, R. G., T. J. R. Harris, and R. A. Lamb. 1984. Fusion proteins of the paramyxovirus simian virus 5: nucleotide sequence of mRNA predicts a highly hydrophobic glycoprotein. Proc. Natl. Acad. Sci. USA 81:6706–6710.
32. Pellett, P. E., K. G. Kousoulas, L. Pereira, and B. Roizman. 1985. Anatomy of herpes simplex virus 1 strain F glycoprotein B gene: primary sequence and predicted protein structure of the wild type and of monoclonal antibody-resistant mutants. J. Virol. 53:243–253.
33. Popovic, M., M. G. Sarugadharan, E. Read, and R. C. Gallo. 1984. A method for detection, isolation and continuous production of cytophathic human T-lymphotropic retroviruses of the HTLV family (HTLV-III) from patients with AIDS and pre-AIDS. Science 224:497–500.
34. Queen, C. L., and L. J. Korn. 1980. Computer analysis of nucleic acids and proteins. Methods Enzymol. 65:595–609.
35. Ratner, L., R. C. Gallo, and F. Wong-Staal. 1985. HTLV-III, LAV, and ARC are variants of the same AIDS virus. Nature (London) 313:636–637.
36. Ratner, L., W. Haseltine, R. Patarca, K. Livak, B. Starcich, S. Josephs, R. E. Doran, J. A. Rafalski, E. A. Whitehorn, K. Baumeister, L. Ivanoff, S. R. Petteway, Jr., I. A. Lautenberger, T. S. Papas, J. Ghrayeb, J. Chang, R. C. Gallo, and F. Wong-Staal. 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature (London) 313:277–284.
37. Robert-Guroff, M., M. Brown, and R. C. Gallo. 1985. HTLV-neutralizing antibodies in AIDS and ARC. Nature (London) 316:72–74.
38. Salinovich, O., S. L. Payne, R. C. Montelaro, K. A. Hussain, C. J. Issel, and K. L. Schnorr. 1986. Rapid emergence of novel antigenic sites and genetic variants of equine infectious anemia virus during persistent infection. J. Virol. 57:71–80.
39. Sanchez-Pescador, R., M. D. Power, J. P. Barr, K. S. Steimer, M. M. Stempien, S. L. Brown-Shimer, W. W. Gee, A. Renard, A. Randolph, J. A. Levy, D. Dina, and P. A. Luciw. 1985. Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2). Science 227:484–492.
40. Shaw, G. M., B. H. Hahn, S. K. Arya, J. E. Groopman, R. C. Gallo, and F. Wong-Staal. 1984. Molecular characterization of human T-cell leukemia (lymphoma) virus type III in the acquired immune deficiency syndrome. Science 226:1165–1171.
41. Shaw, G. M., M. E. Harper, B. H. Hahn, L. G. Epstein, C. D. Gaidusek, R. W. Price, B. A. Navia, C. K. Petito, C. J. O'Hara, E.-S. Cho, J. M. Oleska, F. Wong-Staal, and R. C. Gallo. 1985. HTLV-III infection in brains of children and adults with AIDS-encephalopathy. Science 227:177–182.
42. Sonigo, P., M. Alizon, K. Staskus, D. Klatzmann, S. Cole, O. Danos, E. Retzel, P. Tiollais, A. Haase, and S. Wain-Hobson. 1985. Nucleotide sequence of the visna lentivirus: relationship to the AIDS virus. Cell 42:369–382.
43. Starcich, B. R., B. H. Hahn, G. M. Shaw, P. D. McNeely, S. Modrow, H. Wolf, E. S. Parks, W. P. Parks, S. F. Josephs, R. C. Gallo, and F. Wong-Staal. 1986. Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS. Cell 45:637–648.
44. Tainer, J. A, E. D. Getzoff, H. Alexander, R. A. Houghten, A. J. Olson, R. A. Lerner, and W. A. Hendrickson. 1984. The reactivity of anti-peptide antibodies is a function of the atomic mobility of sites in a protein. Nature (London) 312:127–134.
45. Wain-Hobson, S., P. Sonigo, O. Danos, S. Cole, and M. Alizon. 1985. Nucleotide sequence of the AIDS virus, LAV. Cell 40:9–17.
46. Weiss, R. A., P. R. Chapham, R. Cheingson-Popov, A. G. Dayleish, C. A. Carne, I. V. D. Weller, and R. S. Tedder. 1985. Neutralizing antibodies to human T-cell lymphotropic virus type III. Nature (London) 316:69–72.
47. Westhof, E., D. Altschuh, D. Moras, A. C. Bloomer, A. Mondragon, A. Klug, and M. H. V. van Regenmortel. 1984. Correlation between segmental mobility and the location of antigenic determinants in proteins. Nature (London) 311: 123–126.
48. Wickner, W. T., and H. F. Lodish. 1985. Multiple mechanisms of protein insertion into and across membranes. Science 230:400–407.
49. Wilbur, W. J., and D. J. Lipman. 1983. Rapid similarity searches of nucleic acid and protein data banks. Proc. Natl. Acad. Sci. USA 80:726–730.
50. Wong-Staal, F., G. M. Shaw, B. H. Hahn, S. Z. Salahuddin, M. Popovic, P. D. Markham, R. Redfield, and R. C. Gallo. 1985. Genomic diversity of human T-lymphotropic virus type III. Science 229:759–762.

HIV Molecular Immunology Database

Immunology DB

DB Help
CTL search
T Helper search
AB search
Epitope Maps
Tools & Links
Home

Immunology Tools

Epilign
PeptGen
Motif Scan
Sequence Locator
ELF

Publications

FAQ
Alignments
Tutorials
Reviews
Compendia
Links

Databases

Sequence DB
Resistance DB
Immunology DB
Vaccine Trials DB

HIV Molecular Immunology Compendium

The 1998 Immunology Compendium in PDF format.

You need Adobe Acrobat to read these documents; it is available for all major platforms.

Part I. HIV CTL Epitopes
How To Use This Section (136 K)

Tables

- p17
- p24
- p15
- Pol Proteins
- Tat
- Rev
- Vif
- gp120
- gp41
- Nef

CTL Epitopes Sorted by HLA Restricting Element

Maps

- p17
- p24 and p15
- Pol Proteins
- Rev and gp120
- gp41
- Nef

Alignments

- p17
- p24
- p15
- Pol Proteins
- Rev
- gp120
- gp41

- Nef

References

Part II. HIV Helper T-Cell Epitopes

- How To Use This Section

Tables

- p17
- p24
- p15
- Vpr
- Pol Proteins
- Tat
- Vif
- Vpu
- Rev
- gp120
- gp41
- Nef

Maps

- p17, p24 and p15
- Vpr and Pol Proteins
- Tat and Vif
- Rev and Vpu
- gp120, gp41 and Nef

References

Part III. HIV Monoclonal Antibodies

How To Use This Section

Tables

- p17
- p24
- p7
- p15
- Protease
- RT
- Integrase
- Tat
- Rev
- Vif
- gp120
- gp41
- Nef

Maps

- p17 and p24
- p7-p6 and p15
- RT, In, Pro
- Tat and Rev
- gp120
- gp41
- Nef

Indices
References

Part IV. Review Articles

Structure of the Core of the HIV-1 gp120 exterior Envelope Glycoprotein
Richard Wyatt, Peter D. Kwong, Wayne A. Hendrickson and Joseph G. Sodroski

SIV and SHIV CTL Epitopes Identified in Macaques
Todd M. Allen and David I. Watkins

Systematic Identification of Optimal HIV-1 CTL Epitopes
Christian Brander and Bruce D. Walker

Numbering Positions in HIV Relative to HXB2CG
Bette T. Korber, Brian T. Foley, Carla L. Kuiken, Satish K. Pillai and Joseph G. Sodroski

MHC Tetramer Analyses of CD8+ T Cell Responses to HIV and SIV
John D. Altman and Jeffrey T. Safrit

Last modified: Wed Feb 13 13:47:18 MST 2002

---

immuno@t10.lanl.gov

Los Alamos NATIONAL LABORATORY

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 265, No. 18, Issue of June 25, pp. 10373-10382, 1990
Printed in U.S.A.

Assignment of Intrachain Disulfide Bonds and Characterization of Potential Glycosylation Sites of the Type 1 Recombinant Human Immunodeficiency Virus Envelope Glycoprotein (gp120) Expressed in Chinese Hamster Ovary Cells*

(Received for publication, January 25, 1990)

Cordelia K. Leonard, Michael W. Spellman, Lavon Riddle‡, Reed J. Harris, James N. Thomas§, and Timothy J. Gregory‡¶

From the Departments of Medicinal and Analytical Chemistry, ‡Recovery Process Research and Development and §Cell Culture Research and Development, Genentech, Inc., South San Francisco, California 94080

This report describes the structural characterization of the recombinant envelope glycoprotein (rgp120) of human immunodeficiency virus type 1 produced by expression in Chinese hamster ovary cells. Enzymatic cleavage of rgp120 and reversed-phase high performance liquid chromatography were used to confirm the primary structure of the protein, to assign intrachain disulfide bonds, and to characterize potential sites for *N*-glycosylation. All of the tryptic peptides identified were consistent with the primary structure predicted from the cDNA sequence. Tryptic mapping studies combined with treatment of isolated peptides with *Staphylococcus aureus* V8 protease or with peptide:*N*-glycosidase F followed by endoproteinase Asp-N permitted the assignment of all nine intrachain disulfide bonds of rgp120. The 24 potential sites for *N*-glycosylation were characterized by determining the susceptibilities of the attached carbohydrate structures to peptide:*N*-glycosidase F and to endo-β-*N*-acetylglucosaminidase H. Tryptic mapping of enzymatically deglycosylated rgp120 was used in conjunction with Edman degradation and fast atom bombardment-mass spectrometry of individually treated peptides to determine which of these sites are glycosylated and what types of structures are present. The results indicate that all 24 sites of gp120 are utilized, including 13 that contain complex-type oligosaccharides as the predominant structures, and 11 that contain primarily high mannose-type and/or hybrid-type oligosaccharide structures.

The major envelope glycoprotein (gp[1]120) of HIV-1 has been the object of intensive investigation since the initial identification of HIV-1 as the etiological agent of acquired immunodeficiency syndrome (Barre-Sinoussi *et al.*, 1983). The gp120 molecule is of interest as a vaccine candidate (Berman *et al.*, 1988; Arthur *et al.*, 1987), as the mediator of viral attachment via the virus receptor CD4 (Dalgleish *et al.*, 1984; Klatzman *et al.*, 1984) and as an agent with immunosuppressive effects of its own (Shalaby *et al.*, 1987; Diamond *et al.*, 1988). It is also a potential mediator of the pathogenesis of HIV-1 in acquired immunodeficiency syndrome (Siliciano *et al.*, 1988; Sodroski *et al.*, 1986). The gp120 molecule is synthesized as part of a membrane-bound glycoprotein, gp160 (Allan *et al.*, 1985). Via a host-cell-mediated process, gp160 is cleaved to form gp120 and the integral membrane protein gp41 (Robey *et al.*, 1985). Together gp120 and gp41 form the spikes observed on the surface of newly released HIV-1 virions (Gelderblom *et al.*, 1987). As there is no covalent attachment between gp120 and gp41, free gp120 is released from the surface of virions and infected cells (Gelderblom *et al.*, 1985).

The gp120 molecule consists of a polypeptide core of 60,000 daltons; extensive modification by *N*-linked glycosylation increases the apparent molecular weight of the molecule to 120,000 (Lasky *et al.*, 1986). The amino acid sequence of gp120 contains five relatively conserved domains interspersed with five hypervariable domains (Modrow *et al.*, 1987; Willey *et al.*, 1986). The hypervariable domains contain extensive amino acid substitutions, insertions, and deletions. Sequence variations in these domains result in up to 25% overall sequence variability between gp120 molecules from the various viral isolates. Despite this variation, several structural and functional elements of gp120 are highly conserved. Among these are the ability of gp120 to bind to the viral receptor CD4, the ability of gp120 to interact with gp41 to induce fusion of the viral and host cell membranes, the positions of the 18 cysteine residues in the gp120 primary sequence, and the positions of 13 of the approximately 24 *N*-linked glycosylation sites in the gp120 sequence.

The disulfide bonding pattern within gp120 and the positions of actual oligosaccharide moieties on the molecule would be useful information for directing mutagenesis and fragmentation studies aimed at defining the functional domains of gp120 and sites for potential pharmacological interuption of its functions (*e.g.* type-common neutralizing epitopes). This information has been difficult to obtain due to the small amounts of gp120 available from natural sources, the complexity of the disulfide bonding and oligosaccharide structures in gp120, and uncertainty regarding the functionality or structural relevance (Moore *et al.*, 1990) of rgp120 produced in non-mammalian systems. We have been able to produce large amounts of two different rgp120 fusion proteins in a mammalian cell system (Lasky *et al.*, 1986). This has allowed us

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed.

[1] The abbreviations used are: gp, glycoprotein; AAA, amino acid analysis; CHO, Chinese hamster ovary; DTT, dithiothreitol; endo H, endo-β-*N*-acetylglucosaminidase H; FAB-MS, fast atom bombardment-mass spectrometry; gD1, herpes simplex type 1 glycoprotein D; HIV, human immunodeficiency virus; HPLC, high performance liquid chromatography; PNGase F, peptide: *N*-glycosidase F; PTH, phenylthiohydantoin; RCM, reduced and *S*-carboxymethylated; rgp, recombinant glycoprotein; SIV, simian immunodeficiency virus; TPCK, L-1-*p*-tosylamido-2-phenylethyl chloromethyl ketone.

to elucidate all nine of the disulfide bonds, the positions of the glycosylation sites that are utilized, and the type of oligosaccharide moiety present at each site in rgp120 from the III_B isolate of HIV-1 produced in CHO cells.

EXPERIMENTAL PROCEDURES

Materials—Recombinant gp120 proteins were produced in CHO cells and purified by immunoaffinity chromatography as previously described (Lasky *et al.*, 1986). DTT, iodoacetic acid, and 2-acetamido-1-β-(L-aspartamido)-1,2-dideoxy-D-glucose (GlcNAc-Asn) were obtained from Sigma. HPLC/Spectro Grade trifluoroacetic acid (Pierce Chemical Co.), Acetonitrile UV (American B&J), and Milli Q™ water (Millipore) were used for reversed-phase HPLC. The enzymes used were TPCK trypsin from Worthington Biomedical Corp., endoproteinase Asp-N ("sequencing grade") obtained from Boehringer Mannheim GmbH, *Staphylococcus aureus* V8 protease from ICN ImmunoBiologicals, and PNGase F (N-Glycanase™) and endo H from Genzyme.

Reduction and S-Carboxymethylation—Recombinant gp120 (2.0 mg of CL44) was dialyzed against 0.36 M Tris buffer, pH 8.6, containing 8 M urea and 3 mM EDTA. DTT was added to a concentration of 10 mM, and the sample was incubated for 4 h at ambient temperature. The sample was then treated with 25 mM iodoacetic acid in the dark for 30 min at ambient temperature. The reaction was quenched with excess DTT, the sample was dialyzed against 0.1 M ammonium bicarbonate, and then lyophilized.

Treatment of RCM rgp120 with PNGase F—RCM rgp120 (0.5 mg) was reconstituted in 0.1 ml of 0.25 M sodium phosphate, pH 8.6, containing 10 mM EDTA and 0.02% NaN_3. PNGase F was added to the sample in the ratio of 12.5 units/mg of protein, and the sample was incubated overnight at 37 °C. RCM rgp120 treated with PNGase F was dialyzed against 0.1 M ammonium bicarbonate.

Treatment of RCM rgp120 with Endo H—RCM rgp120 (0.5 mg) was reconstituted in 0.1 ml of 0.05 M sodium phosphate, pH 6.0, containing 0.02% NaN_3. Endo H (2 units/ml) was added to the sample in the ratio of 0.1 unit/mg of protein and the sample was incubated overnight at 37 °C. RCM rgp120 treated with endo H was dialyzed against 0.1 M ammonium bicarbonate.

Treatment with TPCK-Trypsin—Samples of untreated, PNGase F-treated, and endo H-treated RCM rgp120 (0.5-mg aliquots of CL44) in 0.1 M ammonium bicarbonate were treated at ambient temperature with TPCK-trypsin by the addition of aliquots of enzyme (enzyme to substrate ratio of 1:100 w/w) at 0 and 6 h of incubation. The digestion was stopped after 24 h by freezing the samples. For disulfide determinations, a sample of rgp120 (0.5 mg of 9AA) was treated with TPCK-trypsin using the same conditions.

Treatment of Tryptic Peptides with PNGase F Followed by Endoproteinase Asp-N—Peptides (ranging from 0.5 nmol to 3.7 nmol) purified by reversed-phase HPLC of a 9AA tryptic digest were reconstituted in 0.05 M sodium phosphate, pH 7.0, containing 0.02% NaN_3 (0.05 ml). PNGase F (5 units in 0.06 ml of 0.05 M sodium phosphate, pH 7.0, containing 0.02% NaN_3) was added, and the samples were incubated for 20 h at 37 °C. Endoproteinase Asp-N (2 μg) was then added, and the samples were incubated for 20 h at 37 °C.

Treatment of Tryptic Peptides with S. aureus V8 Protease—Peptides (3.0 nmol) purified by reversed-phase HPLC of a 9AA tryptic digest were reconstituted in 0.05 M sodium phosphate, pH 7.0, containing 0.02% NaN_3 (0.04 ml). V8 protease (5 μg) was added at 0 and 7 h, and the sample was incubated for 24 h at 37 °C.

Treatment of CL44 Peptides with Endo H Followed by PNGase F—Peptides (typically 3 nmol) purified by reversed-phase HPLC were reconstituted in 0.05 M sodium phosphate, pH 6.0, containing 0.02% NaN_3 (0.1 ml). Endo H (0.05 unit in 0.025 ml of 0.05 M sodium phosphate, pH 6.0, containing 0.02% NaN_3) was added, and the sample was incubated for 20 h at 37 °C. PNGase F (6.25 units) and 0.5 M sodium phosphate, pH 10.3, containing 0.02 M EDTA and 0.02% NaN_3 (0.125 ml) were then added, and the sample was incubated for 20 h at 37 °C.

Reversed-phase HPLC—Tryptic digests were fractionated by reversed-phase HPLC on a 5-μm Vydac C_{18} endcapped column (4.6 × 250 mm). After equilibration with 0.1% aqueous trifluoroacetic acid, the elution of tryptic peptides was carried out at 1 ml/min with a linear gradient from 0 to 45% acetonitrile containing 0.08% trifluoroacetic acid in 90 min. The system used was a Waters gradient liquid chromatograph consisting of two 6000A pumps, a 720 controller, and a WISP 710B injector, and a Perkin-Elmer LC75 single wavelength UV detector set at 214 nm.

Peptides subjected to further manipulations were fractionated by reversed-phase HPLC on a Vydac C_{18} column (2.1 × 250 mm) equilibrated in 0.1% aqueous trifluoroacetic acid at a flow rate of 0.2 ml/min and a temperature of 40 °C. These peptides were eluted with a linear gradient from 0 to 60% acetonitrile (containing 0.08% trifluoroacetic acid) in 60 min. The system used was a Hewlett-Packard 1090M liquid chromatograph.

Peptide Identification—Peptides collected from reversed-phase HPLC were identified by AAA and/or N-terminal sequence analysis. Samples for AAA were treated with constant boiling HCl at 110 °C *in vacuo* for either 24 or 72 h, depending upon extent of glycosylation. The extended hydrolysis degrades glucosamine, which would otherwise interfere with quantitation of Ile and Leu. Analysis was performed on a Beckman model 6300 amino acid analyzer with ninhydrin detection.

N-terminal sequence analysis was performed on an Applied Biosystems model 477A/120A. The acetonitrile concentration in the equilibration buffer of the PTH analysis system was decreased from 10 to 9% to resolve the PTH derivative of GlcNAc-Asn from DTT.

FAB-MS—FAB mass spectra were acquired on a JEOL HX110HF/HX110HF tandem mass spectrometer operated in a normal two-sector mode. FAB-MS was performed with 6 keV xenon atoms (10 mA emission current). Data were acquired over a mass range of 380–4000 atomic mass units.

RESULTS

Lasky *et al.* (1986) expressed gp120 in CHO cells as a fusion protein using the signal peptide of the herpes simplex gD1. Two such fusion proteins were used in this study. The recombinant glycoprotein used in most of this study (CL44) was expressed as a 498-amino acid fusion protein containing the first 27 residues of gD1 fused to residues 31–501 of gp120 (Lasky *et al.*, 1986). This construction lacks the first cysteine residue of mature gp120. Disulfide assignments were carried out on another recombinant fusion protein (9AA) which contains the first 9 residues of gD1 fused to residues 4–501 of gp120. This restores the first cysteine residue, Cys-24. Carboxyl-terminal analysis of CL44 using carboxypeptidase digestions indicated that glutamic acid residue 479 is the carboxyl terminus of the fully processed molecule secreted by CHO cells (data not shown). The amino acid sequences of these two constructions are given in Fig. 1.

RCM CL44 Tryptic Map—Reversed-phase HPLC tryptic mapping was used to confirm the primary structure of the molecule, to assign intrachain disulfide bonds, and to characterize potential sites for *N*-glycosylation. In experiments not intended to give information about disulfides, the protein was RCM prior to digestion with trypsin. This treatment unfolds the protein and disrupts disulfide bonds, thereby resulting in smaller tryptic fragments than would be obtained with the native molecule.

The reversed-phase HPLC tryptic map of RCM CL44 is shown in Fig. 2. Tryptic peptides were separated by reversed-phase HPLC using an acetonitrile/water system with trifluoroacetic acid as the ionic modifier. As will be discussed below, much of the peak heterogeneity derives from the extremely high (approximately 50% of total mass) carbohydrate content of the molecule. Peaks were collected and subjected to AAA for identification (Table I). In some cases, N-terminal sequence analysis was used for confirmation (these peaks are indicated in Table I). The peaks not assigned a label in Fig. 2 were not identified.

All of the peptides identified were consistent with the primary structure predicted from the cDNA sequence. Of the 38 predicted peptides with three or more amino acids, 36 were identified in the tryptic map of RCM CL44. In addition, four predicted peptides consisting of two amino acids each were also identified (H3, H4, T23, and T35). The tripeptide com-

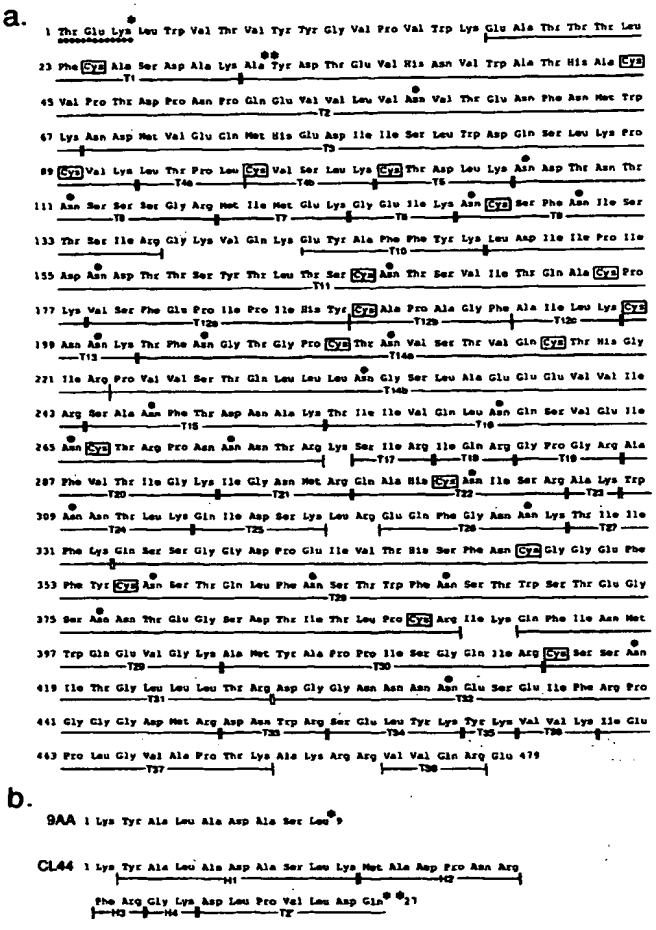


FIG. 1. **Amino acid sequences of the mature envelope glycoprotein (gp120) from the III_B isolate of HIV-1 (*a*) and the N-terminal sequence portion of the recombinant fusion glycoproteins (9AA or CL44) from the herpes simplex gD1 (*b*).** Fusion sites between the gD1 and gp120 segments in the 9AA and CL44 constructions are marked with (*) and (**), respectively. The *letter T* refers to observed tryptic cleavage of the gp120 segment, and the peptides are ordered sequentially starting at the N terminus of the molecule. *Lower case letters* following the T number indicate other unexpected proteolytic cleavages. The *letter H* refers to the observed tryptic cleavage of the herpes simplex gD1 protein portion of CL44. Peptide T2′ contains the fusion site in CL44. The cysteine residues of gp120 are *enclosed in boxes*, and potential *N*-glycosylation sites are indicated with a *dot* above the corresponding asparagine residue.

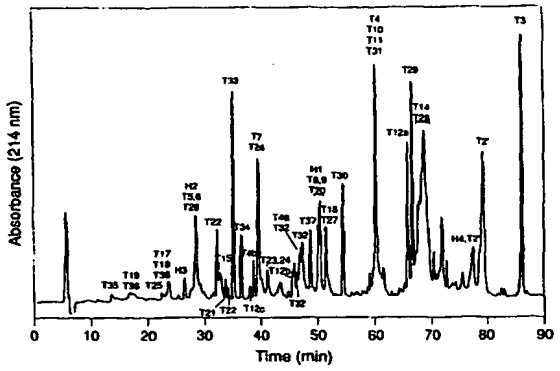


FIG. 2. **Reversed-phase HPLC tryptic map of RCM CL44.** This chromatogram was generated with 7.5 nmol of trypsin-digested RCM CL44. Chromatography conditions were as described under "Experimental Procedures." Peaks were collected and identified by AAA and in some cases confirmed by N-terminal sequence analysis (Table I). Identified peaks are labeled according to the nomenclature given in Fig. 1. Peptides containing potential tryptic sites that were not hydrolyzed are designated by two T numbers separated by a comma.

posed of residues 139–141 (VQK) was not identified in the map and was not given a label in Fig. 1. The only other peptide not identified was T13 (CNNK). Asparagine residue 200 of peptide T13 is a potential glycosylation site and the peptide lacks hydrophobic amino acids. Therefore, this glycopeptide is likely to be extremely hydrophilic and poorly resolved from the salt fraction on the reversed-phase column.

Tryptic cleavage did not occur between peptides T5 and T6 and between peptides T8 and T9. These are designated in Fig. 2 as two T-numbers separated by a comma (T5,6 and T8,9). The absence of cleavage was confirmed by N-terminal sequence analysis of the peptides. In both of these cases, the asparagine residue to the C-terminal side of the cleavage site is a potential *N*-glycosylation site, and it is likely that the carbohydrate moiety interferes with the action of trypsin. Incomplete tryptic cleavage was also observed between peptides H4 and T2′ and between peptides T23 and T24 (H4,T2′ and T23,24).

Several peptides arising from non-tryptic cleavages were observed in the tryptic map of RCM CL44. Two of the predicted tryptic peptides were further cleaved by "chymotrypsin-like" cleavages. Peptide T12 was completely cleaved after tyrosine residue 187 and phenylalanine residue 193 to yield peptides T12a, T12b, and T12c. Peptide T4 was partially hydrolyzed after leucine residue 95 to yield peptides T4a and T4b. Intact peptide T4 was also present.

One of the tryptic peptides, T22 (QAHCNISR) eluted at two different positions (32.4 and 34.1 min) in the RCM CL44 tryptic map. Deglycosylation studies (discussed below) with PNGase F and endo H indicated that the different retention times of the two forms of peptide T22 are not due to carbohydrate differences. It is possible that this retention time heterogeneity results from partial conversion of the N-terminal glutamine residue to pyroglutamic acid (Sanger and Thompson, 1953).

Disulfide Assignments in gp120—Mature gp120 contains 18 cysteine residues (*enclosed in boxes* in Fig. 1) and therefore could contain nine intrachain disulfide bonds. The CL44 construction lacks Cys-24, the first cysteine residue of gp120 (Lasky *et al.*, 1986); therefore, a different construction (*9AA*), in which the first cysteine residue was restored, was purified to approximately the same degree as CL44.[2] Ellman's reagent (Ellman, 1959) was used to demonstrate the absence of free sulfhydral groups in 9AA (data not shown). Therefore, disulfide assignments were determined for the 9AA construction.

Tryptic mapping studies performed without reduction and *S*-carboxymethylation of cysteine residues allowed partial assignment of disulfides. The tryptic map of 9AA is shown in Fig. 3. Peaks were identified by N-terminal sequence analysis (Table *II*). These identifications allowed unequivocal assignment of three of the nine disulfide bonds: between Cys-101 and Cys-127 (peak A, Table II), between Cys-266 and Cys-301 (peak B, Table II), and between Cys-24 and Cys-44 (peak E, Table II).

Peptides containing the remaining cysteine residues were also identified (Table II). Peptide T28 contains 3 cysteine residues and coelutes with peptide T31, which contains 1

[2] L. Riddle, T. Gregory, and D. Dowbenko, unpublished data.

TABLE I

Amino acid compositions of peptides from the tryptic map of RCM CL44

Peptides are listed according to elution position in the tryptic map of RCM CL44 (Fig. 2). Theoretical values derived for peptides labeled in Fig. 1 are in parenthesis. Non-integral theoretical values are the result of coeluting peptides.

Peptide(s)	T35	T19 T36	T25	T17 T18 T38	H3	H2 T5,6 T26[a]	T22	T15	T21	T33	T34
CyA	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMC	0.2	0.0	0.0	0.0	0.0	0.9(1)	0.9(1)	0.1	0.1	0.0	0.0
Asx	0.4	0.2	1.2(1)	0.4	0.0	7.7(9)	1.6(1)	2.9(3)	1.1(1)	1.8(2)	0.1
Thr	0.0	0.0	0.1	0.0	0.0	2.3(3)	0.2	0.9(1)	0.0	0.0	0.0
Ser	0.0	0.0	0.6(1)	0.6(1)	0.0	1.5(3)	0.6(1)	0.5(1)	0.0	0.0	0.6(1)
Glx	0.1	0.0	0.9(1)	2.0(2)	0.0	2.1(2)	1.0(1)	0.1	0.0	0.1	1.0(1)
Pro	0.0	1.0(1)	0.0	0.0	0.0	0.9(1)	0.0	0.1	0.0	0.0	0.0
Gly	0.1	2.1(2)	0.1	0.3	0.1	2.5(2)	0.0	0.1	0.9(1)	0.1	0.1
Ala	0.1	0.0	0.1	0.0	0.0	1.0(1)	1.4(1)	1.9(2)	0.1	0.0	0.0
Cys	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0
Val	0.2	0.7(2)	0.1	1.2(2)	0.1	0.1	0.0	0.1	0.0	0.1	0.1
Met	0.0	0.1	0.0	0.0	0.0	0.7(1)	0.0	0.0	0.9(1)	0.0	0.0
Ile	0.3	0.0	0.9(1)	2.2(2)	0.0	0.2	1.0(1)	0.1	1.0(1)	0.0	0.0
Leu	0.1	0.0	0.0	0.0	0.0	1.0(1)	0.0	0.0	0.0	0.0	1.0(1)
Tyr	0.0(1)	0.4	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	1.0(1)
Phe	0.0	0.0	0.0	0.0	1.0(1)	1.0(1)	0.2	0.9(1)	0.1	0.0	0.0
His	0.0	0.0	0.0	0.0	0.0	0.0	1.0(1)	0.1	0.1	0.0	0.0
Lys	1.2(1)	0.0(1)	0.9(1)	0.5	0.0	2.0(2)	0.2	1.2(1)	0.0	0.0	1.0(1)
Trp	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0(1)	0.0
Arg	0.1	1.1(1)	0.1	2.6(3)	1.0(1)	3.6(2)	1.0(1)	0.1	1.0(1)	1.1(1)	0.0

Peptide(s)	T12c	T4b	T7 T24[a]	T23,24	T12b	T32	T4a T32[a]	T37	H1 T9,8 T20[a]	T18 T27[a]
CyA	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMC	0.0	0.0(1)	0.2	0.2	0.9(1)	0.3	0.1	0.1	0.2(0.2)	0.3(0.3)
Asx	0.1	0.3	1.0(2)	2.0(2)	0.3	5.2(6)	1.6(1.8)	0.3	0.0(0.7)	1.5(1.5)
Thr	0.0	0.3	0.9(1)	1.3(1)	0.2	0.6	0.7(0.5)	0.9(1)	1.0(1.1)	1.5(1.8)
Ser	0.1	0.0(1)	0.2	0.3	0.1	0.0(1)	0.3(0.3)	0.3	0.5(0.7)	0.2(0.3)
Glx	0.2	0.6	1.1(1)	0.6	0.1	2.0(2)	0.6(0.5)	1.1(1)	0.3(0.2)	0.9(0.9)
Pro	0.1	0.3	0.2	0.2	0.9(1)	1.0(1)	0.0(0.0)	1.9(2)	0.2	0.3(0.3)
Gly	0.2	0.5	0.3	0.4	1.1(1)	4.7(5)	1.4(1.5)	1.0(1)	1.0(1.1)	0.1
Ala	1.0(1)	0.0	0.3	0.9(1)	1.0(2)	0.1	0.0	0.9(1)	1.6(1.8)	0.0
Cys	0.0	0.1	0.0	0.2	0.1	0.0	0.0	0.0	0.0	0.0
Val	0.3	1.2(1)	0.2	0.5	0.1	0.4	0.1	0.9(1)	0.9(0.9)	0.6(0.6)
Met	0.0	0.0	1.7(2)	0.0	0.0	0.7(1)	0.2(0.2)	0.0	0.0	0.0
Ile	1.0(1)	0.3	1.0(1)	0.5	0.1	1.1(1)	0.4(0.3)	1.2(1)	1.4(1.5)	2.4(2.7)
Leu	1.0(1)	0.9(1)	1.0(1)	0.9(1)	0.0	0.1	1.0(1.0)	0.9(1)	0.6(0.6)	0.3(0.3)
Tyr	0.0	0.0	0.1	0.2	0.0	0.1	0.0	0.1	0.3(0.3)	0.0
Phe	0.0	0.0	0.0	0.5	1.1(1)	1.2(1)	0.3(0.3)	0.2	1.0(1.1)	0.9(0.9)
His	0.0	0.3	0.1	0.2	0.1	0.2	0.1	0.0	0.0	0.0
Lys	1.0(1)	0.9(1)	1.9(2)	1.7(2)	0.0	0.1	0.0	1.0(1)	1.3(1.4)	0.8(0.9)
Trp	0.0	0.0	0.0(1)	0.0(1)	0.0	0.0	0.0	0.0	0.0	0.0
Arg	0.2	0.0	0.0	0.2	0.1	1.8(2)	0.5(0.6)	0.2	0.2(0.2)	0.6(0.6)

Peptide(s)	T30	T4 T10 T11 T31[a]	T12a	T29	T14 T28[a]	H4,T2'	T2'	T3
CyA	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMC	0.1	2.1(2.5)	0.1	0.2	2.0(2.1)	0.5(1)	0.9(1)	1.1(1)
Asx	0.2	4.0(4.3)	0.7	1.3(1)	5.2(4.9)	8.5(9)	8.6(9)	3.9(4)
Thr	0.2	4.7(5.1)	0.4	0.5	4.6(5.8)	3.5(4)	3.3(4)	0.0
Ser	0.6(1)	2.1(4.0)	0.7(1)	0.3	3.4(5.4)	0.6	0.2	1.3(2)
Glx	1.1(1)	1.7(1.5)	1.2(1)	2.7(3)	4.5(4.5)	5.5(5)	5.1(5)	4.0(4)
Pro	2.0(2)	1.6(1.7)	1.9(2)	0.2	1.6(1.6)	3.9(4)	4.0(4)	1.1(1)
Gly	1.9(1)	0.9(0.8)	0.3	1.1(1)	4.1(4.2)	4.5(1)	0.2	0.1
Ala	1.9(2)	1.5(1.5)	0.1	0.1	0.4(0.3)	2.1(2)	2.0(2)	0.1
Cys	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.0
Val	0.2	1.1(1.0)	1.1(1)	1.0(1)	1.8(2.3)	8.0(8)	8.0(8)	2.1(2)
Met	0.8(1)	0.0	0.0	0.7(1)	0.0	0.6(1)	0.9(1)	1.6(2)
Ile	1.9(2)	3.2(3.6)	1.9(2)	0.9(1)	1.5(1.6)	0.3	0.2	1.9(2)
Leu	0.1	4.3(4.7)	0.3	0.3	2.1(2.2)	3.2(3)	3.1(3)	2.0(2)
Tyr	1.0(1)	2.2(2.3)	0.8(1)	0.1	0.6(0.5)	1.4(1)	0.9(1)	0.0
Phe	0.1	1.7(1.6)	1.0(1)	1.0(1)	3.9(2.8)	1.7(1)	1.0(1)	0.1
His	0.0	0.0	1.0(1)	0.1	0.7(0.8)	2.2(2)	2.0(2)	1.0(1)
Lys	0.1	1.8(1.8)	0.1	0.0(1)	0.1	1.7(2)	1.1(1)	2.0(2)
Trp	0.0	0.0	0.0	0.0(1)	0.0(1.0)	0.0(2)	0.0(2)	0.0(1)
Arg	1.0(1)	0.0(0.8)	0.1	0.1	1.0(1.1)	0.0	0.1	0.0

[a] Confirmed by N-terminal sequence analysis.

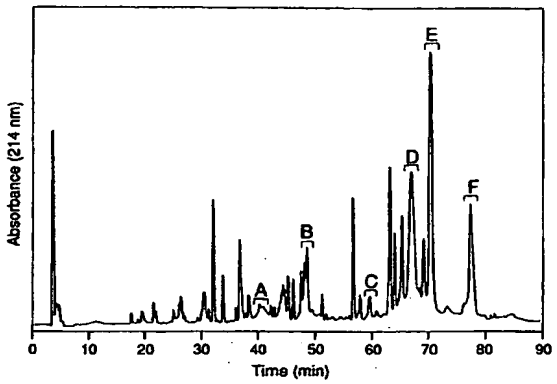


FIG. 3. **Reversed-phase HPLC tryptic map of 9AA.** This chromatogram was generated with 6.8 nmol of sample. Chromatography conditions were as described under "Experimental Procedures." Peaks containing cysteine residues were identified by N-terminal sequence analysis. These identifications are summarized in Table II.

TABLE II

Identification of cysteine-containing peptides from the tryptic map of 9AA

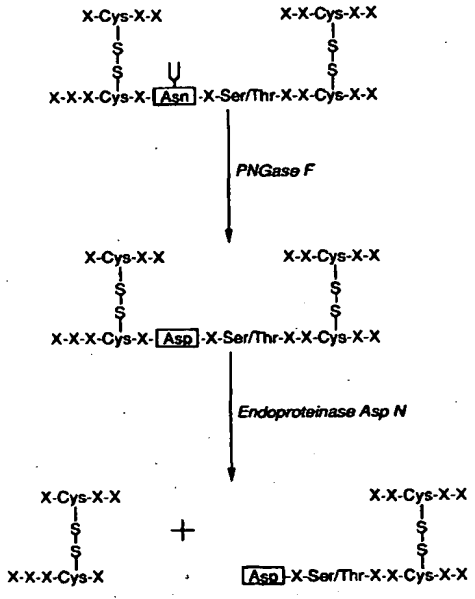
Cys-containing peaks from the tryptic map of 9AA (Fig. 3) were identified by N-terminal sequence analysis. Cysteines in boxes joined by a solid line represent disulfide bond assignments. Cysteines in boxes joined by dotted lines represent disulfide bonds that could not be assigned unambiguously in this experiment. Partial cleavages are indicated by a parenthesis. Cysteines are labeled by an amino acid number and peptides are labeled with T-numbers corresponding to the nomenclature used in Fig. 1.

Peak	Cys-Containing Peptides	Recovery (nmol)
A	101 (T5,6) [C]TDLKNDTNTNSSSGR (GEIK)N[C]SFNISTSIR 127 (T9,8)	2.3
B	266 (T10) TIIVQLNQSVEIN[C]TRPNNNTR QAH[C]NISR 301 (T22)	4.5
C	(T12a,b) 188 VSFEPIPIHY[C]APAGF 198 (T13) [C]NNK (T14a) TFNGTGP[C]TNVSTVQ[C]THGIR 209 217	0.6
D	(T28) 348 355 QSSGGDPEIVTHSFN[C]GGEFFY[C]NSTQLFNSTWFNSTWSTE- -GSNNTEGSDTITLP[C]R (T31) 388 [C]SSNITGLLLTR 415	3.8
E	24 (T1) EATTTLF[C]ASDAK AYDTEVHNVWATHA[C]VPTDPNPQEVVLVNVTENFNMWK 44 (T2)	3.9
	(T12) 188 VSFEPIPIHY[C]APAGFAILK 198 (T13) [C]NNK (T14) TFNGTGP[C]TNVSTVQ[C]THGIRPVVSTQLLLNGSLAEEEVVIR 209 217	2.1
F	(T3) 89 NDMVEQMHEDIISLWDQSLKP[C]VK 96 (T4) LTPL[C]VSLK (T11) LDIIPIDNDTTSYTLTS[C]NTSVITQA[C]PK 166 175	4.6

cysteine residue (peak D, Table II). Peptide T11 contains 2 cysteine residues and coelutes with peptides T3 and T4, each of which contains a single cysteine residue (peak F, Table II). Similarly, peptide T14 contains 2 cysteine residues and coelutes with peptides T12 and T13, each of which has a single cysteine residue (peaks C and E, Table II). In each of these cases more than one disulfide bond was present in the group of tryptic peptides, thereby preventing unambiguous assignment. These tryptic peptides were further manipulated as described below to introduce selective cleavage between cysteine residues located on a single peptide.

The procedure used to cleave between the cysteine residues of peptides T11 and T14 is summarized (Scheme 1). Each of the peptides has a potential *N*-linked glycosylation site located between the cysteine residues. The peptides were treated with PNGase F, which removes asparagine-linked carbohydrate while converting the attachment asparagine residue to aspartic acid (Tarentino *et al.*, 1985). The resulting aspartic acid residue serves as a point for selective cleavage of the peptides with endoproteinase Asp-N (Drapeau, 1980). The peptides were separated by reversed-phase HPLC and identified by N-terminal sequence analysis.

The HPLC chromatogram obtained after treatment of peptides T12, T13, and T14 (*peak C*, Fig. 3) with PNGase F followed by endoproteinase Asp-N is given in Fig. 4*a*, and the



SCHEME 1

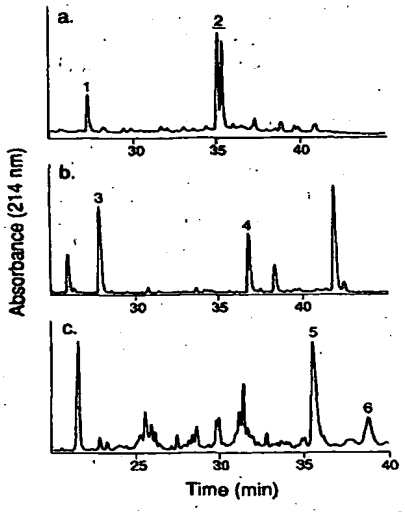


FIG. 4. **Further manipulations of tryptic peptides from the map of 9AA to isolate individual disulfides.** The chromatograms are details of microbore reversed-phase HPLC separations of peptides resulting from treatment of peptides T12, T13, and T14 (*peak C*, Fig. 3) with PNGase F followed by endoproteinase Asp-N (*a*), treatment of peptides T3, T4, and T11 (*peak F*, Fig. 3) with PNGase F followed by endoproteinase Asp-N (*b*), and treatment of peptides T28 and T31 (*peak D*, Fig. 3) with *S. aureus* V8 protease (*c*). Chromatography conditions were as described under "Experimental Procedures." Peak identifications were determined by N-terminal sequence analysis and are given in Table III.

sequences of relevant peptides are given in Table III. The results indicate that rgp120 has disulfide bonds between Cys-198 and Cys-209 and between Cys-188 and Cys-217 (Table III). Comparable manipulation of *peak E* gave similar results. Treatment of peptides T3, T4, and T11 (*peak F*, Fig. 3) with PNGase F followed by endoproteinase Asp-N allowed the recovery of fragments that demonstrated the presence of disulfide bonds between Cys-89 and Cys-175 and between Cys-96 and Cys-166 (Fig. 4*b* and Table III).

The last two disulfide bonds were assigned by treating

TABLE III

Assignment of disulfides from peptides isolated in Fig. 4

The tryptic peptides that could not be assigned unambiguously in Table II were further manipulated as described in Fig. 4. Peaks were identified by N-terminal sequence analysis.

Peak	Sequence
1	188 C N DGTGP C T 209
2	188 EPIPIHY C APAGF DVSTVD C THG(IR) 217
3	89 DQSLKP C VK DTSVITQA C PK 175
4	96 LTPL C VSLK DDTTSYTLTS C 166
5	348 IVTHSFN C GGE C SSNITGLLLTR 415
6	355 FFY C NSTQLFNSTWFNSTWSTE TITLP C R 388

peptides T28 and T31 (*peak D*, Fig. 3) with V8 protease to cleave to the carboxy side of the glutamic acid and aspartic acid residues (Drapeau *et al.*, 1972) located between the cysteine residues of T28. The chromatogram obtained after V8 protease digestion of T28 and T31 is given in Fig. 4*c* and the sequences of the relevant peptides are given in Table III. The results demonstrated the presence of disulfide bonds between Cys-348 and Cys-415 and between Cys-355 and Cys-388.

Thus, the combined results of the tryptic mapping analysis and the further selective degradations permitted the assignment of all nine intrachain disulfide bonds of rgp120. Parallel experiments performed on CL44 produced similar results for the eight disulfide bonds remaining in that construction (not shown). The disulfide bond assignments of rgp120 are summarized in Fig. 6.

Glycosylation Sites of gp120—Mature gp120 contains 24 potential sites for *N*-glycosylation, as recognized by the sequence Asn-Xaa-Ser(Thr) (Kornfeld and Kornfeld, 1985). These sites are indicated by a *dot* above the corresponding asparagine residue in Fig. 1. In the present study, tryptic mapping of enzymatically deglycosylated CL44 was used in conjunction with Edman degradation and FAB-MS of individually treated peptides to determine which of the 24 potential *N*-glycosylation sites are glycosylated and which contain less fully processed (*i.e.* high mannose-type or hybrid-type) oligosaccharides.

The two enzymes used for deglycosylation were PNGase F and endo H. PNGase F releases all types of *N*-linked oligosaccharide structures by cleavage of the β-aspartylglucosylamine linkage (Tarentino *et al.*, 1985). Endo H releases only high mannose-type and hybrid-type oligosaccharide structures by cleaving between the two core *N*-acetylglucosamine residues (Tai *et al.*, 1977). Deglycosylation of a peptide can be monitored by the increase in retention time of the peak corresponding to the glycopeptide in the reversed-phase elution profile. Thus, it was possible to determine which peptides were glycosylated by treatment with PNGase F and, on the basis of susceptibility to endo H, to distinguish those with attached high mannose-type and/or hybrid-type oligosaccharides as the predominant structures.

The 24 potential glycosylation sites of CL44 are contained in 14 tryptic glycopeptides. Thirteen of these glycopeptides were identified in the tryptic map of RCM CL44 (Fig. 2). As mentioned above, T13 (CNNK) was not identified. The tryptic maps of PNGase F-treated RCM CL44 and endo H-treated RCM CL44 are compared with the RCM CL44 tryptic map in Fig. 5. The peaks corresponding to glycopeptides are labeled in each of the three tryptic maps.

As would be expected for a heavily glycosylated molecule, treatment of RCM CL44 with PNGase F (Fig. 5*b*) simplified the tryptic map significantly. Typically, the peaks corresponding to potential glycopeptides in the RCM CL44 tryptic map (Fig. 5*a*) were broad and often appeared as multiplets. Deglycosylation resulted in sharp, single peaks for each peptide, indicating that the glycopeptide peak multiplicity and broadness was due to carbohydrate heterogeneity.

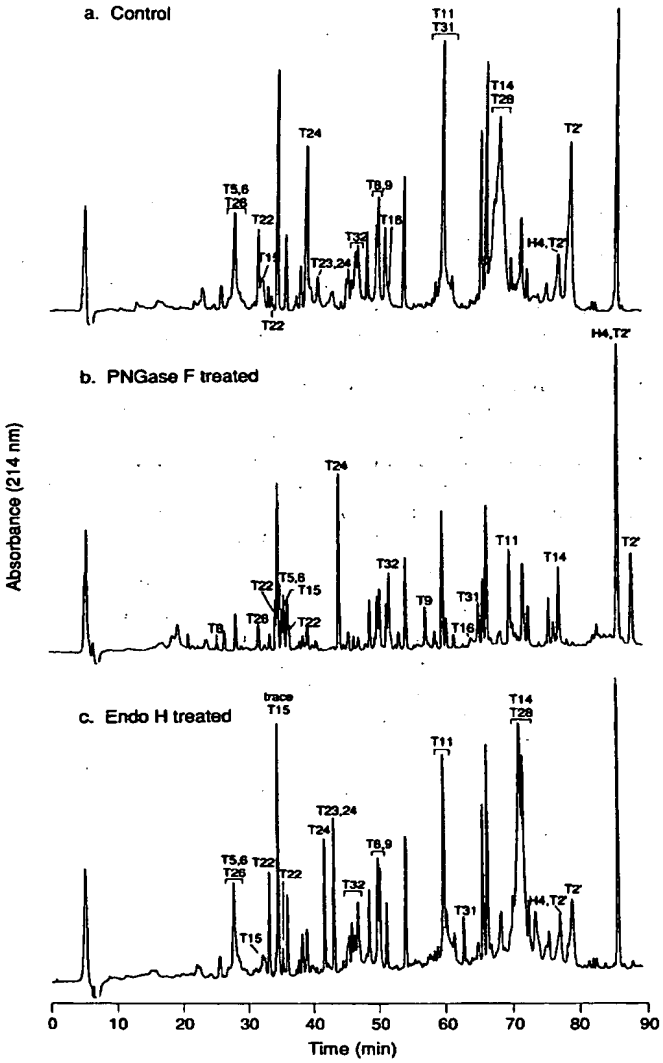


FIG. 5. **Reverse-phase HPLC tryptic maps of endoglycosidase-treated RCM CL44.** The chromatograms are tryptic maps of untreated RCM CL44 (*a*), PNGase F-treated RCM CL44 (*b*), and endo H-treated RCM CL44 (*c*). Each tryptic map was generated with 7.5 nmol of sample. Chromatography conditions were as described under "Experimental Procedures." Peaks were collected and identified by *AAA* (data not shown). Glycopeptide peaks are labeled according to the nomenclature in Fig. 1.

Of the 13 potential glycopeptides that had been identified in the tryptic map of RCM CL44, all were shifted to later retention times in the tryptic map of PNGase F-treated material. This demonstrates that at least 13 of the 24 potential sites are glycosylated. Peptide T28 was not recovered after deglycosylation. This peptide contains a large number of nonpolar amino acids and, after removal of the hydrophilic carbohydrate moieties, may bind irreversibly to the HPLC column. As described above, peptide T22 elutes at two positions in the RCM CL44 tryptic map presumably as a result of conversion of the N-terminal glutamine to pyroglutamic acid. The retention times of both of the T22 peaks were altered in the deglycosylated material produced by treatment with both PNGase F and endo H, confirming that the difference between these forms of peptide T22 in the RCM CL44 tryptic map was not due to carbohydrate heterogeneity.

The tryptic map of endo H-treated RCM CL44 (Fig. 5*c*) indicated that six of the 13 tryptic glycopeptides were endo H-susceptible (peptides T14, T16, T22, T24, T28, and T31). In addition, a small amount of peptide T15 showed endo H susceptibility. For each of these glycopeptides, the elution time of the endo H-treated glycopeptide was earlier than that of the corresponding PNGase F-treated glycopeptide. This is due to the hydrophilic *N*-acetylglucosamine residue that remains attached to the asparagine residue following endo H treatment. Peptide T16 was not identified in the tryptic map of endo H-treated RCM CL44. This peptide contains three potential glycosylation sites and was poorly recovered under any circumstances.

Conclusions as to the type of glycosylation present on each of the tryptic glycopeptides based on susceptibility to PNGase F and endo H are summarized in Table IV. Seven of the 13 glycopeptides identified in the tryptic map of RCM CL44 contain only a single glycosylation site and thus could be characterized unambiguously with regard to enzyme susceptibility. Peptides T2′ (Asn-58), T26 (Asn-326), and T32 (Asn-433) were deglycosylated only by PNGase F and, therefore, contain attached complex-type oligosaccharide structures. Peptides T22 (Asn-302), T24 (Asn-309), and T31 (Asn-418) were susceptible to both PNGase F and endo H and, therefore,

TABLE IV

Assignment of glycosylation type to RCM CL44 tryptic peptides by susceptibility to PNGase F and endo H

Susceptibility to PNGase F or endo H was determined by an increase in the retention time of a peptide in the tryptic map of RCM CL44 (Fig. 5). PNGase F releases all types of *N*-linked oligosaccharide structures, whereas endo H releases only high mannose and hybrid oligosaccharide structures.

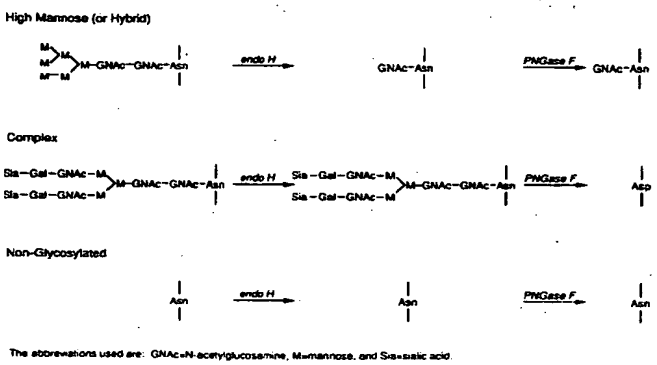
Tryptic Peptide [a]	*Glycosylation Sites (Asn Residue #)*	*Susceptible To PNGase F*	*Susceptible To Endo H*	*Glycosylation Type*
T2′	58	Yes	No	Complex
T5	106,111	Yes	No	Complex [b]
T9	126,130	Yes	No	Complex [b]
T11	156,167	Yes	No	Complex [b]
T14	204,211,232	Yes	Yes	High Mannose, Hybrid, and/or Complex [c]
T15	246	Yes	Trace	Complex (Trace High Mannose and/or Hybrid)
T16	259,265,271	Yes	Yes	High Mannose, Hybrid, and/or Complex [c]
T22	302	Yes	Yes	High Mannose and/or Hybrid
T24	309	Yes	Yes	High Mannose and/or Hybrid
T26	326	Yes	No	Complex
T28	356,362,367,376	Yes	Yes	High Mannose, Hybrid, and/or Complex [c]
T31	418	Yes	Yes	High Mannose and/or Hybrid
T32	433	Yes	No	Complex

[a] T13 not found.
[b] Either or both sites glycosylated.
[c] Endo H susceptible glycosylation at one or more site(s).

carry high mannose-type and/or hybrid-type oligosaccharide structures. Peptide T15 is only partially susceptible to endo H; therefore, Asn-246 carries primarily complex-type oligosaccharides but must also have some attached high mannose-type and/or hybrid-type oligosaccharide structures.

Peptides T6, T9, and T11 each contain two potential glycosylation sites. Each peptide was deglycosylated by PNGase F but not by endo H indicating the presence of mostly complex-type oligosaccharide structures. In order to determine whether one or both of the potential glycosylation sites in each peptide were actually glycosylated, the PNGase F-treated glycopeptides were subjected to either FAB-MS or Edman degradation. Treatment with PNGase F converts the attachment asparagine residue to aspartic acid during deglycosylation (Tarentino *et al.*, 1985). This conversion can be detected by FAB-MS as an increase of 1 atomic mass unit in the mass of the peptide for each site deglycosylated (Carr and Roberts, 1986) or by Edman degradation by the appearance of the PTH derivative of aspartic acid at the appropriate cycles. FAB-MS of deglycosylated peptide T5,6 revealed an ion corresponding to the peptide mass plus 2 atomic mass units ($[MH]^+$ observed: m/z 1772.6; calculated: m/z 1772.7). FAB-MS of deglycosylated peptide T9 gave similar results ($[MH]^+$ observed: m/z 1301.8; calculated: m/z 1301.5). Edman degradation was performed instead of FAB-MS on deglycosylated peptide T11 because of its high molecular weight (>2000 a.m.u.). Aspartic acid was observed in cycles 8 (derived from Asn-156) and 19 (derived from Asn-167). These combined results indicate the presence of complex-type oligosaccharide structures attached to Asn residues 106, 111, 126, 130, 156, and 167.

The remaining three glycopeptides identified in the tryptic map of RCM CL44 contained multiple potential glycosylation sites and were endo H susceptible. Peptides T14, T16, and T28 account for a total of 10 potential glycosylation sites. Characterization of each glycosylation site was achieved by Edman degradation of HPLC-purified peptides that had been subjected to treatment with endo H followed by PNGase F (Scheme 2). When endo H releases the high mannose-type and hybrid-type oligosaccharide structures, it leaves an *N*-acetylglucosamine residue attached to the asparagine residue of the peptide (Tarentino *et al.*, 1974). PNGase F will not remove this *N*-acetylglucosamine residue but will release the remaining *N*-linked oligosaccharide structures by cleavage of the β-aspartylglucosylamine bond, resulting in conversion of the attachment asparagine residue to aspartic acid (Chu, 1986). Therefore, treatment with endo H followed by PNGase F will yield asparagine at an unglycosylated site, GlcNAc-Asn at a glycosylation site that contained primarily high mannose-type and/or hybrid-type oligosaccharide structures, and aspartic acid at a glycosylation site that carried primarily complex-type oligosaccharide structures. Paxton *et al.* (1987) have shown that it is possible to detect the PTH derivative of GlcNAc-Asn after Edman degradation. Using this approach, it was possible to characterize the remainder of the glycosylation sites of CL44. For example, treatment of glycopeptide T16, which contains three potential *N*-glycosylation sites, with endo H followed by PNGase F resulted in the appearance of the PTH derivative of GlcNAc-Asn at cycles 7 and 13 and the appearance of PTH-Asp at cycle 19 during Edman degradation. Thus, glycopeptide T16 carries primarily high mannose-type and/or hybrid-type oligosaccharides at Asn-259 and Asn-265 and complex-type oligosaccharides at Asn-271. The results of these experiments are summarized in Table V and indicate that CL44 contains complex-type oligosaccharide



SCHEME 2

TABLE V

Assignment of glycosylation type to RCM CL44 tryptic glycopeptides containing multiple potential glycosylation sites

Characterization of multiple potential glycosylation sites on RCM CL44 tryptic glycopeptides was achieved by Edman degradation of HPLC-purified peptides subjected to treatment with endo H followed by PNGase F. Edman degradation of deglycosylated peptides shows either an Asn residue at an unglycosylated site, a GlcNAc-Asn at a glycosylation site to which had been attached high mannose or hybrid oligosaccharide structures, or an Asp residue at a glycosylation site which had carried complex type oligosaccharide structures.

Tryptic Peptide	*Asn Residue #*	*Residue Observed*	*Glycosylation Type*
T14	204	GlcNAc-Asn	High Mannose and/or Hybrid
	211	GlcNAc-Asn	High Mannose and/or Hybrid
	232	GlcNAc-Asn	High Mannose and/or Hybrid
T16	259	GlcNAc-Asn	High Mannose and/or Hybrid
	265	GlcNAc-Asn	High Mannose and/or Hybrid
	271	Asp	Complex
T28	356	GlcNAc-Asn	High Mannose and/or Hybrid
	362	GlcNAc-Asn	High Mannose and/or Hybrid
	367	Asp	Complex
	376	Asp	Complex

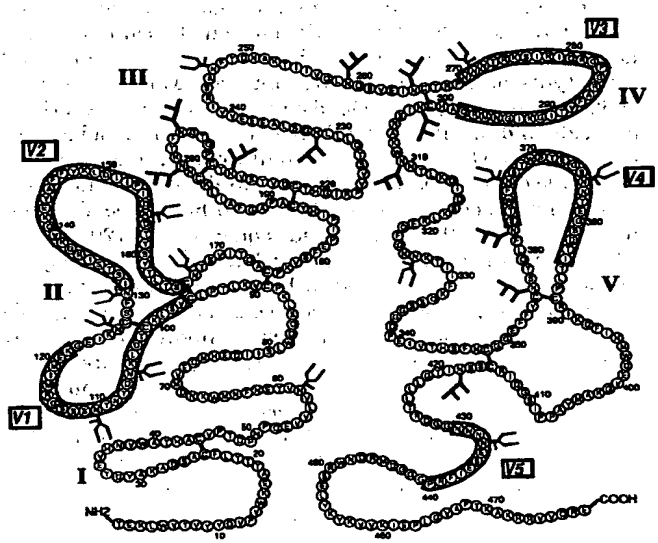


FIG. 6. **Schematic representation of gp120 showing disulfides and glycosylation sites.** Glycosylation sites containing high mannose-type and/or hybrid-type oligosaccharide structures are indicated by a (Y) and glycosylation sites containing complex-type oligosaccharide structures are indicated by a (Ψ). *Roman numerals* label the five disulfide-bonded domains. The five hypervariable regions of Modrow *et al.* (1987) are enclosed in *boxes* and labeled *V1-V5*.

structures at Asn residues 271, 367, and 376 and high mannose-type and/or hybrid-type oligosaccharide structures at Asn residues 204, 211, 232, 259, 265, 356, and 362.

Peptide T13, which contains the remaining glycosylation site, was not identified in any of the tryptic maps presented in this paper. However, FAB-MS data obtained from the void peak of a tryptic map of RCM CL44 treated with endo H followed by PNGase F revealed an ion corresponding to MH^+ for that peptide containing an attached *N*-acetylglucosamine residue (observed: *m/z* 740.1; calculated: *m/z* 740.4). The presence of peptide T13 in the void peak was further confirmed by AAA. Therefore, we conclude that Asn-200 is glycosylated and carries primarily high mannose-type and/or hybrid-type oligosaccharide structures.

The data presented here demonstrate that all 24 potential glycosylation sites of gp120 are utilized, that 13 sites contain primarily complex-type oligosaccharide structures while 11 sites contain primarily high mannose-type and/or hybrid-type oligosaccharide structures. The type of glycosylation at each site is summarized in Fig. 6.

DISCUSSION

We have determined the disulfide bonding pattern and the attachment positions of oligosaccharide moieties of rgp120 from the III_B isolate of HIV-1. A schematic representation of this information is presented in Fig. 6. The rgp120 molecules from which the structural data were obtained possess the functional properties attributed to gp120 produced by HIV-1 virions including high affinity CD4 binding (Lasky *et al.*, 1987), and HIV-1 neutralizing antigenicity (Lasky *et al.*, 1986). We therefore conclude that the CHO-expressed gp120 is properly folded and that the disulfide-bonded domains reported here for the recombinant molecules are representative of those occurring in gp120 produced by HIV-1 virions.

Functional Aspects of gp120 Structure—The gp120 molecule comprises five disulfide-bonded loop structures. The first and fourth are simple loops formed by single disulfide bonds while the second, third, and fifth are more complex arrays of loops formed by nested disulfide bonds. The fourth disulfide-bonded domain (residues 266–301) has been shown to contain significant type-specific neutralizing epitopes (Matsushita *et al.*, 1988; Rusche *et al.*, 1988; Goudsmit *et al.*, 1988; Javaherian *et al.*, 1989) and the fifth disulfide-bonded domain (residues 348–415) has been shown to be important for CD4 binding (Lasky *et al.*, 1987; Kowalski *et al.*, 1987). No direct functional correlates have been described for the other three disulfide-bonded domains. The amino acid sequence of gp120 varies to a large extent between different viral isolates but the majority of the variability is localized in hypervariable regions which punctuate the otherwise relatively conserved sequences (Willey *et al.*, 1986; Modrow *et al.*, 1987). Modrow *et al.* (1987) have identified five hypervariable regions which are characterized by sequence variation, insertions, and deletions. Four of these hypervariable regions correspond to well-delineated loops as indicated in Fig. 6. With the exception of the third hypervariable loop (disulfide-bonded domain IV) the functional significance of these regions is unknown.

The positions of the cysteine residues and, presumably, the disulfide bonding pattern in gp120 are highly conserved between isolates. Among HIV-1 isolates, the only exception to this conservation is the Z3 isolate (Willey *et al.*, 1986) which has an additional pair of cysteine residues in the fourth hypervariable domain (residues 363–384). These residues most likely form a tenth disulfide bond in the gp120 from this isolate. The presence of this extra bond in such a hypervariable region probably has no more effect on the structure and function of the molecule than the other sequence variations that occur in that region. In HIV-2 and SIV, the positions of the cysteine residues in disulfide-bonded domains I, II, IV, and V are conserved (Myeres *et al.*, 1989). In domain III there are two additional pairs of cysteine residues (three in SIV isolate MM142) which are presumed to be disulfide bonded within a finger-like domain III structure analagous to that illustrated in Fig. 6. Another major difference between HIV-1, HIV-2, and SIV is that hypervariable region V2 is reduced to five amino acids in HIV-2 and SIV. The functional significance of the differences between HIV-1, HIV-2, and SIV is unknown at this time.

One of the most important functions of gp120 is its ability to bind to CD4 and thereby mediate the attachment of virions to susceptible cells (Klatzman *et al.*, 1984; Dalgleish *et al.*, 1984). The CD4-binding function has been localized by mutagenesis and structural studies (Lasky *et al.*, 1987; Kowalski *et al.*, 1987) to the region between residues 320 and 450, which includes the fifth disulfide-bonded domain. Lasky *et al.* (1987) showed that deletion of residues 396 to 407 and mutagenesis of Ala-402 to Asp abolished CD4 binding. They also mapped the epitope of a monoclonal antibody that blocks gp120-CD4 binding to residues 392–402. Kowalski *et al.* (1987) identified three regions as being involved with CD4 binding. Insertions between residues 333–334, 388–390, and 442–443 abolished CD4 binding. In addition, a deletion of residues 441–479 abolished CD4 binding while deletion of residues 362–369 within the fourth hypervariable region had no effect on binding. Cordonnier *et al.* (1989) have shown that mutagenesis of Trp-397 to Tyr or Phe decreases CD4 binding and changes to Ser, Gly, Val, or Arg abolish binding. Nygren *et al.* (1988) have reported that a proteolytic fragment of gp120 from residue 322 to near the C terminus retains the ability to bind to CD4. The results of these studies indicate that the CD4 binding capacity of gp120 is localized to the region between residues 320 and 450 and more specifically to the residues around 333–334, 442–443, and the sequence between 388 and 407.

In the course of efforts to map the epitope of monoclonal antibody 5C2-E5 which blocks gp120-CD4 binding, Lasky *et al.* (1987) treated rgp120 (CL44) with acetic acid to cleave the protein at aspartic acid residues (Ingram, 1963) and isolated the peptide fragment 383–426 from a column of immobilized anti-gp120 monoclonal antibody 5C2-E5. Digestion of reduced rgp120 yielded the same fragment. Consequently, it was concluded that a disulfide bond existed between Cys residues 388 and 415. In the analysis reported here we have failed to find this disulfide bond and, instead, have consistently found the disulfide bonds between Cys-355 and Cys-388, and between Cys-348 and Cys-415 as summarized in Fig. 6. We believe that the true disulfide-bond assignment is as indicated in Fig. 6 and that the acetic acid digestion produced some disulfide bond rearrangement (Ryle and Sanger, 1955) in the earlier work.

The Oligosaccharides of gp120—Approximately 50% of the apparent molecular mass of gp120 is carbohydrate. The structures of the oligosaccharide moieties released by hydrazinolysis of CL44 rgp120 have been exhaustively analyzed (Mizuochi *et al.*, 1988a; Mizuochi *et al.*, 1988b). These authors found that 33% of the *N*-linked oligosaccharides were of the high mannose type, 4% were of the hybrid type, and 63% were of the complex type. Of the complex oligosaccharides 90% were fucosylated and 94% were sialylated. The complex structures were approximately 4% monoantennary, 61% biantennary, 19% triantennary, and 16% tetraantennary. No *O*-linked oligosaccharides were found. Geyer *et al.* (1988) have

analyzed the oligosaccharides of gp120 from the III_B isolate of HIV-1-infected human cells. They found that high mannose-type oligosaccharides accounted for approximately 50% of the carbohydrate structures. The remaining structures were fucosylated, partially sialylated bi-, tri-, and tetraantennary complex-type oligosaccharides. No novel carbohydrate structures, or moieties that would be expected to act as heterophile antigens in man, have been isolated from gp120 from either source.

We have shown here that all 24 glycosylation sites are utilized, and that 13 of the 24 sites contain complex-type oligosaccharides as the predominant structures while 11 contain primarily hybrid and/or high mannose structures. The demonstration of endo H-susceptible structures at 11 of the 24 sites is consistent with the earlier results of Mizuochi *et al.* (1988a, 1988b) who determined that nearly 40% of the total oligosaccharide structures released from rgp120 were hybrid and/or high mannose-type oligosaccharides.

The 24 potential *N*-linked glycosylation sites in the gp120 sequence are conserved to a large extent between different viral isolates (Willey *et al.*, 1986; Modrow *et al.*, 1987). Based on the gp120 sequence comparisons in these references, 13 of the sites on gp120 from the III_B isolate of HIV-1 are absolutely conserved; these include eight of the 11 sites that carry predominantly hybrid-type and/or high mannose-type oligosaccharides. Thus, the less fully processed (*i.e.* endo H-susceptible) oligosaccharides of gp120 are found preferentially at the most conserved glycosylation sites. The remaining sites (eight complex and three hybrid/high mannose) are relatively conserved, even though many of them occur in the hypervariable regions. The positions of these sites may shift or be deleted, but there is always one or more new site(s) within 5–10 residues of the reference III_B site. Studies by Willey *et al.* (1988) demonstrated that mutagenesis of Asn-232 to Gln decreased the infectivity of virions containing the mutant gp120 molecules without affecting CD4 binding or syncitium formation. At this time, no particular functional significance can be attributed to the type of oligosaccharide structure at any of the sites.

The role of the carbohydrate moieties on gp120 in CD4 binding has been investigated by several authors (Lifson *et al.*, 1986; Matthews *et al.*, 1987; Fenouillet *et al.*, 1989). Those that employed enzymatic deglycosylation in the presence of detergents (Lifson *et al.*, 1986; Matthews *et al.*, 1987) have concluded that the carbohydrates are not directly involved with the binding but that they are required to maintain the conformation of gp120 necessary for binding. In contrast, Fenouillet *et al.* (1989) enzymatically deglycosylated gp120 without detergent and demonstrated that the CD4 binding affinity was preserved. It therefore appears that the carbohydrate moieties of gp120 are not required for its binding to CD4 but that the conformational stability of gp120 to detergents is lost after deglycosylation.

The rgp120 used for these determinations is functionally and structurally equivalent to gp120 produced by HIV-1-infected cells. The structural data presented here will be useful in future attempts to manipulate the structure of gp120 in order to better understand the biology of the virus and to produce an effective vaccine.

Acknowledgments—We wish to thank Dr. Phil Berman and Dr. Larry Lasky for many helpful discussions, Dr. John Chakel for FAB-MS analysis, and Carol Morita and Anne Stone for preparation of illustrations.

REFERENCES

Allan, J. S., Coligan, J. E., Barin, F., McLane, M. F., Sodroski, J. G., Rosen, C. A., Haseltine, W. A., Lee, T. H., and Essex, M. (1985) *Science* **228**, 1091–1094

Arthur, L. O., Pyle, S. W., Nara, P. L., Bess, J. W., Jr., Gonda, M. A., Kelliher, J. C., Gilden, R. V., Robey, W. G., Bolognesi, D. P., Gallo, R. C., and Fischinger, P. J. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 8583–8587

Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W., and Montagnier, L. (1983) *Science* **220**, 868–871

Berman, P. W., Groopman, J. E., Gregory, T. J., Weiss, R. A., Clapham, P. R., Ferriani, R., Riddle, L., Shimasaki, C., Lucas, C., Lasky, L. A., and Eichberg, J. W. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 5200–5204

Carr, S. A., and Roberts, G. D. (1986) *Anal. Biochem.* **157**, 396–406

Chu, F. K. (1986) *J. Biol. Chem.* **261**, 172–177

Cordonnier, A., Montagnier, L., and Emerman, M. (1989) *Nature* **340**, 571–574

Dalgleish, A. G., Beverely, P. C. L., Clapham, P. R., Crawford, D. H., Greaves, M. F., and Weiss, R. A. (1984) *Nature* **312**, 763–767

Diamond, D. C., Sleckman, B. P., Gregory, T., Lasky, L. A., Greenstein, J. L., and Burakoff, S. J. (1988) *J. Immunol.* **141**, 3715–3717

Drapeau, G. R. (1980) *J. Biol. Chem.* **255**, 839–840

Drapeau, G. R., Boily, Y., and Houmard, J. (1972) *J. Biol. Chem.* **247**, 6720–6726

Ellman, G. L. (1959) *Arch. Biochem. Biophys.* **82**, 70–77

Fenouillet, E., Clerget-Raslain, B., Gluckman, J. C., Guetard, D., Montagnier, L., and Bahraoui, E. (1989) *J. Exp. Med.* **169**, 807–822

Gelderblom, H. R., Reupke, H., and Pauli, G. (1985) *Lancet* **ii**, 1016–1017

Gelderblom, H. R., Hausmann, E. H. S., Ozel, M., Pauli, G., and Koch, M. A. (1987) *Virology* **156**, 171–176

Geyer, H., Holschbach, C., Hunsmann, G., and Schneider, J. (1988) *J. Biol. Chem.* **263**, 11760–11767

Goudsmit, J., Boucher, C. A. B., Meloen, R. H., Epstein, L. G., Smit, L., van der Hoek, L., and Bakker, M. (1988) *AIDS* **2**, 157–164

Ingram, V. (1963) *Methods Enzymol.* **6**, 831–848

Javaherian, K., Langlois, A. J., McDanal, C., Ross, K., Eckler, L. I., Jellis, C. L., Profy, A. T., Rusche, J. R., Bolognesi, D. P., Putney, S. D., and Matthews, T. J. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 6768–6772

Klatzmann, D., Champagne, E., Chamaret, S., Gruest, J., Guetard, D., Hercend, T., Gluckman, J.-C., and Montagnier, L. (1984) *Nature* **312**, 767–768

Kornfeld, R., and Kornfeld, S. (1985) *Annu. Rev. Biochem.* **54**, 631–664

Kowalski, M., Potz, J., Basiripour, L., Dorfman, T., Goh, W. C., Terwilliger, E., Dayton, A., Rosen, C., Haseltine, W., and Sodroski, J. (1987) *Science* **237**, 1351–1355

Lasky, L. A., Groopman, J. E., Fennie, C. W., Benz, P. M., Capon, D. J., Dowbenko, D. J., Nakamura, G. R., Nunes, W. M., Renz, M. E., and Berman, P. W. (1986) *Science* **233**, 209–212

Lasky, L. A., Nakamura, G., Smith, D. H., Fennie, C., Shimasaki, C., Patzer, E., Berman, P., Gregory, T., and Capon, D. J. (1987) *Cell* **50**, 975–985

Lifson, J., Coutre, S., Huang, E., and Engleman, E. (1986) *J. Exp. Med.* **164**, 2101–2106

Matsushita, S., Robert-Guroff, M., Rusche, J., Koito, A., Hattori, T., Hoshino, H., Javaherian, K., Takatsuki, K., and Putney, S. (1988) *J. Virol.* **62**, 2107–2114

Matthews, T. J., Weinhold, K. J., Lyerly, H. K., Langlois, A. J., Wigzell, H., and Bolognesi, D. P. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 5424–5428

Mizuochi, T., Spellman, M. W., Larkin, M., Solomon, J., Basa, L. J., and Feizi, T. (1988a) *Biochem. J.* **254**, 599–603

Mizuochi, T., Spellman, M. W., Larkin, M., Solomon, J., Basa, L. J., and Feizi, T. (1988b) *Biomed. Chromatogr.* **2**, 260–270

Modrow, S., Hahn, B. H., Shaw, G. M., Gallo, R. C., Wong-Stahl, F., and Wolf, H. (1987) *J. Virol.* **61**, 570–578

Moore, J. P., McKeating, J. A., Jones, I. M., Stephens, P. E., Clements, G., Thomson, S., and Weiss, R. A. (1990) *AIDS* **4**, 307–315

Myers, G., Rabson, A., Josephs, S., Smith, T., Berzofsky, J., and Wong-Stahl, F. (eds) (1989) *Human Retroviruses and AIDS, LA-UR*, 89–743, Los Alamos National Laboratory, Los Alamos

Nygren, A., Bergman, T., Matthews, T., Jornvall, H., and Wigzell, H. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 6543–6546

Paxton, R. J., Mooser, G., Pande, H., Lee, T. D., and Shively, J. E. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 920–924

Robey, W. G., Safai, B., Oroszlan, S., Arthur, L. O., Gonda, M. A.,

Gallo, R. C., and Fischinger, P. J. (1985) *Science* **228**, 593–595
Rusche, J. R., Javaherian, K., McDanal, C., Petro, J., Lynn, D. L., Grimaila, R., Langlois, A., Gallo, R. C., Arthur, L. O., Fischinger, P. J., Bolognesi, D. P., Putney, S. D., and Matthews, T. J. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 3198–3202
Ryle, A. P., and Sanger, F. (1955) *Biochem. J.* **60**, 535–540
Sanger, F., and Thompson, E. O. P. (1953) *Biochem. J.* **53**, 366–374
Shalaby, M. R., Krowka, J. F., Gregory, T. J., Hirabayashi, S. E., McCabe, S. M., Kaufman, D. S., Stites, D. P., and Ammann, A. J. (1987) *Cell. Immunol.* **110**, 140–148
Siliciano, R. F., Lawton, T., Knall, C., Karr, R. W., Berman, P., Gregory, T., and Reinherz, E. L. (1988) *Cell* **54**, 561–575
Sodroski, J., Goh, W. C., Rosen, C., Campbell, K., and Haseltine, W. A. (1986) *Nature* **322**, 470–474
Tai, T., Yamashita, K., and Kobata, A. (1977) *Biochem. Biophys. Res. Commun.* **78**, 434–441
Tarentino, A. L., Plummer, T. H., Jr., and Maley, F. (1974) *J. Biol. Chem.* **249**, 818–824
Tarentino, A. L., Gomez, C.M., and Plummer, T.H., Jr. (1985) *Biochemistry* **24**, 4665–4671.
Willey, R. L., Rutledge, R. A., Dias, S., Folks, T., Theodore, T., Buckler, C. E., and Martin, M. A. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 5038–5042
Willey, R. L., Smith, D. H., Lasky, L. A., Theodore, T. S., Earl, P. L., Moss, B., Capon, D. J., and Martin, M. A. (1988) *J. Virol.* **62**, 139–147

CHAPTER
13

Human Immunodeficiency Virus

Emilio A. Emini
Scott D. Putney

As of January 1990, over 115,000 cases of the acquired immunodeficiency syndrome (AIDS) were recorded in the USA and more than 50% of those afflicted with the disease have died. About 50,000 cases of AIDS are projected to be reported in 1990 (Berkelman and Curran 1989; Centers for Disease Control 1990a, 1990b). The syndrome is characterized by a severe immunodeficiency, manifested by the acquisition of opportunistic infections and the development of rare neoplasias. The syndrome also exhibits signs of central nervous system impairment and, in certain cases, dementia. The disease inevitably results in the death of the patient, usually within several years after its initial clinical appearance.

13.1 AIDS EPIDEMIOLOGY

AIDS is the end-stage clinical result of immune system abnormalities that occur as a consequence of infection with the human immunodeficiency virus type 1 (HIV-1). The virus, a member of the *Lentivirinae* subfamily of the retroviruses, is transmitted by blood and through intimate sexual contact. Upon entering the host, the virus predominantly infects cells that bear the differentiation antigen CD4 on their surface (Dalgleish et al. 1984; Klatz-

mann et al. 1984). This molecule serves as the viral receptor permitting the attachment of the virus to the cell. The most important CD4-bearing cells are helper T-lymphocytes, as well as monocytes and macrophages. Both cell types are essential to the appropriate function of the immune system. Infection with the virus is characterized by a short, acute viremic phase followed by a long period of latency and persistent infection that may last a number of years. The persistent phase usually exhibits a progressive decline of CD4-positive, helper T-lymphocytes, culminating in the eventual loss of immune system function and clinical AIDS.

Several epidemiological patterns of HIV-1 infection have been noted. The first, found predominantly in the developed world, is characterized by the presence of the infection among certain high-risk groups. These groups have traditionally been at risk of acquiring blood-transmitted infections and include male homosexuals and intravenous drug abusers. Heterosexual transmission occurs among the sexual partners of infected drug abusers and mother-to-infant transmission is seen in this population. The second pattern is characteristic of central Africa and seems to be emerging in some South American countries. Here, the infection is widespread in the population, particularly in urban areas with an equal distribution of the infection among males and females, suggesting significant transmission by heterosexual contact. Finally, a third epidemiological pattern has been described in most of the rest of the world in which cases of HIV-1 infection are confined to small groups typically composed of travelers, and their contacts, from Pattern 1 and Pattern 2 areas.

13.2 HIV-1 BIOLOGY

The lentiviruses have been recognized for many years as pathogens of veterinary importance. They include the caprine arthritis-encephalitis virus of goats, the visna/maedi virus of sheep, and the equine infectious anemia virus of horses. In recent years, a number of lentiviruses related to HIV-1 have been identified, most of which mediate the development of immunodeficiency in their respective hosts. These are the feline immunodeficiency virus of cats, the bovine immunodeficiency virus of cows, and many simian immunodeficiency viruses that have been isolated from diverse African and Asian monkey species (Daniel et al. 1985; Gonda et al. 1988; Murphy-Corb et al. 1986; Pedersen et al. 1987). A relative of HIV-1, termed HIV-2, has been isolated from humans in western Africa (Clavel et al. 1986). This virus is more closely related to some of the simian viruses than it is to HIV-1. It has been associated with the development of AIDS in infected humans, but as yet is not as epidemiologically important as HIV-1.

HIV-1 is an enveloped virus that contains a dense nucleoid core structure. The surface of the virion is composed of multiple copies of the envelope

glycoprotein, gp120, a highly glycosylated 120-kDa protein (Allan et al. 1985). The gp120 is noncovalently attached to a transmembrane anchor gp41 glycoprotein. gp120 and gp41 are derived by cleavage during virus maturation from a common gp160 precursor. The virion core is composed of a p17 myristyolated matrix protein located immediately under the virion's envelope and a p24 core protein that packages two copies of the virus' genomic RNA. Both p17 and p24 are derived by cleavage, via a viral protease, from a p55 precursor (Kohl et al. 1988; Muesing et al. 1985; Ratner et al. 1985; Wain-Hobson et al. 1985). The viral core also contains virus-coded reverse transcriptase and integrase enzymes.

Following adsorption of the virus onto its CD4 receptor, the virus' envelope fuses with the plasma membrane of the host cell. It is hypothesized that this fusion is mediated by the hydrophobic amino terminus of gp41 (Kowalski et al. 1987). The viral core enters the cell's cytoplasm and the viral RNA is reverse transcribed to ultimately form the double-stranded DNA provirus that integrates randomly into the cell's genome. Expression of the provirus is tightly controlled by a complex set of viral regulatory proteins (reviewed by Pavlakis and Felber 1990). In infected lymphocytes, high-level virus expression usually occurs when the cell is activated by specific antigens or by mitogens. Virus expression culminates in the production of progeny virions by budding at the cell's plasma membrane. Such high-level virus production typically leads to the destruction of the infected lymphocyte (reviewed by Fauci 1988). This specific cellular destruction may be the primary means of lymphocyte depletion in the host.

Infection of the host cell may also occur by fusion with a virus-infected cell. As with infection by cell-free virus, the fusion is mediated by the initial interaction of viral gp120 on the surface of the virus-expressing cell with CD4 receptors on uninfected cells. The importance of cellular transmission of the virus with respect to transmission of the infection and pathogenesis remains unclear.

Finally, an aspect of the virus' biology, which is particularly pertinent to the development of a vaccine, is the extraordinary amino acid sequence variability exhibited by HIV-1. As with other lentiviruses, HIV-1 exhibits considerable antigenic variation (Hahn et al. 1986). The variability is most prominent in the envelope glycoprotein that can vary by as much as 15% among different virus isolates. There exists a pattern of constant and variable regions throughout the glycoprotein. The conserved regions may participate in functions common to all isolates such as CD4-receptor binding. The variable regions result from the activity of a particularly unfaithful reverse transcriptase (Ricchetti and Buc 1990) and there may be immune selective pressures that result in the establishment of distinct virus variants (e.g., neutralizing antibody resistant mutants). This phenomenon has been noted with the equine infectious anemia virus (Payne et al. 1987) and the consequences of this for HIV-1 will be considered later.

13.3 ANTI-HIV-1 IMMUNE RESPONSES

The immune responses to HIV-1 that occur after infection have been studied in both infected humans and in chimpanzees infected with known variants of the virus. As with most viral infections, the initial antibody response is usually seen after the acute viremia that occurs within weeks following infection. Antibodies develop to all of the viral structural proteins, including the core and envelope antigens as well as the reverse transcriptase and integrase enzymes.

Antibodies that neutralize virus infectivity in vitro first appear a few weeks after infection. This initial virus-neutralizing activity is usually specific for the infecting virus variant and the antigenic determinant that binds this neutralizing antibody has been mapped to a specific region of the gp120 envelope glycoprotein (see Section 13.4.3.1). Subsequently, a neutralizing activity develops which is broader in its in vitro neutralizing capacity. This common group activity has not been specifically mapped as yet, but appears to be directed to determinants expressed by the viral gp120 and/or gp41 glycoproteins (Berkower et al. 1989; Chanh et al. 1986; Dalgleish et al. 1988; Evans et al. 1989; Goudsmit et al. 1988c; Haigwood et al. 1990; Ho et al. 1987, 1988; Nara et al. 1987; Zagury et al. 1988).

Antibodies that mediate antibody-dependent cellular cytotoxicity (ADCC) also develop early in infection (Ljunggren et al. 1987). In addition, the establishment of a persistent HIV-1 infection results in the induction of a persistent cytotoxic T-lymphocyte (CTL) response directed against virus-infected cells. This response is mediated predominantly by CD8-positive T-cells and is major histocompatibility complex (MHC) class I-restricted. Most of the viral proteins expressed by infected cells have been defined as targets of the virus-specific CTLs. These include the viral envelope and core proteins as well as the reverse transcriptase (Hosmalin et al. 1990; Koenig et al. 1988; McChesney et al. 1990; Nixon et al. 1988; Plata et al. 1987; Riviere et al. 1989; Shepp et al. 1988; Takahashi et al. 1988; Walker et al. 1987, 1988, 1989).

13.4 HIV-1 VACCINES: BASIC STUDIES

13.4.1 The Vaccine's Goal

Traditionally, the goal of a successful vaccine has been the establishment of an immune state that can prevent or attenuate the development of clinical disease following infection with the pathogen. Vaccine-induced immune responses typically are not effective in totally preventing infection by the pathogen. However, an immune state that does not prevent the establishment of the HIV-1 persistent infection will probably not be effective in preventing clinical AIDS. As discussed above, the immune responses against the virus following infection are substantial and of a broad spectrum. Yet, the progression to immune deficiency and clinical disease is inexorable. Hence, a successful AIDS vaccine must elicit an immune state that will prevent the establishment of the HIV-1 persistent/latent infection.

The biology of HIV-1 during the initial phase of infection is not well understood and the efficiency with which the virus develops a state of persistent infection is not known. It is difficult to predict which of the many immune responses directed against the virus or virus-infected cells would most likely be effective in preventing infection. If one assumes that the virus can establish its persistent infection almost immediately after introduction into the host, then an immune response directed against the virus, instead of one directed against virus-infected cells, would theoretically be more likely to prevent the persistent infection.

Given this, early efforts to develop an HIV-1 vaccine have focused on the elicitation of immune responses directed against cell-free virus. The elicitation of immune responses against virus-infected cells has not been ignored, but this is a decidedly more difficult and less clear goal.

Finally, it should be noted that immune responses that might be effective in preventing infection by the parenteral route might not be effective in preventing infection by sexually transmitted virus. However, our understanding regarding this issue is limited and requires much additional study.

13.4.2 Animal Models

HIV-1 can be used to experimentally infect chimpanzees when inoculated intravenously (Alter et al. 1984). To date, this remains the most realistic model for testing the efficacy of prototype HIV-1 vaccines, especially since the immune response of chimpanzees to experimental immunogens is very similar, if not identical, to that of humans. Unlike humans, no chimpanzee has yet to develop clinical AIDS. However, this does not limit use of the chimpanzee as an infection model.

HIV-1 has also been demonstrated to inefficiently infect rabbits (Kulaga et al. 1988, 1989; Filice et al. 1988) and to infect scid mice that have been reconstituted with human lymphoid tissue or primary lymphocytes (Namikawa et al. 1988; Mosier et al. 1988). Both models may prove useful for testing the efficacy of various immune responses by passive transfer. The use of simian immunodeficiency virus (SIV)-infected monkeys in vaccine studies will be discussed in Section 13.6.

13.4.3 The HIV-1 Principal Neutralization Determinant (PND)

13.4.3.1 Definition of the PND. The most logical immune response directed against the virus that would be effective in preventing infection in the host would be an antibody capable of neutralizing the infectivity of the

virus. Such antibody activities appear during infection as discussed in Section 13.3. The earlier, type-specific activity has been extensively studied and, as will be noted, is the primary focus of at least some of the HIV-1 vaccine research efforts. The neutralizing antibody response that develops later, although of broader specificity, is less potent in its activity than the type-specific response, and has been poorly and inconsistently elicited by a few test immunogens (Chanh et al. 1986; Dalgleish et al. 1988; Evans et al. 1989; Ho et al. 1987, 1988; Zagury et al. 1988).

The HIV-1 gp120 envelope glycoprotein, whether purified from virus-infected cells or from various recombinantly engineered systems, was found to elicit low levels of type-specific virus-neutralizing antibodies when inoculated into test animals (Arthur et al. 1987; Berman et al. 1988; Hu et al. 1987; Lasky et al. 1986; Matthews et al. 1986; Nara et al. 1988; Putney et al. 1986; Robey et al. 1986; Rusche et al. 1987). The determinant to which these antibodies bind was mapped to a specific fragment of gp120 defined by amino acid residues 300–330 (Goudsmit et al. 1988a, 1988b; Ho et al. 1987; Javaherian et al. 1989; Kenealy et al. 1989; Rusche et al. 1988). Synthetic peptide versions of this region of gp120 were shown to elicit virus-neutralizing antibody responses and a number of gp120-specific neutralizing monoclonal antibodies were directed specifically to this region (Fung et al. 1987; Ho et al. 1987; Kenealy et al. 1989; Matsushita et al. 1988; Palker et al. 1988; Skinner et al. 1988a, 1988b; Thomas et al. 1988).

The region is located in the third hypervariable, V3, domain of gp120 and is known to exist as a disulfide closed loop on the glycoprotein (Leonard et al. 1990; Modrow et al. 1987) (Figure 13–1). It has been termed the HIV-1 principal neutralization determinant (PND) and its amino acid sequence variability among different isolates of the virus accounts for the type specificity of the antibody directed against it. Nonetheless, in spite of this specificity, anti-PND antibody is strongly neutralizing in vitro, is capable of neutralizing the virus even after the virus has bound to its CD4 receptor, and can prevent the fusion of virus-infected and uninfected cells (Linsley et al. 1988; Nara 1989; Skinner et al. 1988a). These properties are highly desirable for an antibody activity that is hoped can establish a complete barrier against HIV-1 infection.

13.4.3.2 Protective Effect of Anti-PND Antibody. Several studies have yielded results which strongly suggest that anti-PND antibody can protect chimpanzees from HIV-1 infection. In the first, Emini et al. (1990a) treated in vitro the prototypic IIIb variant of HIV-1 independently with four different purified immunoglobulin G (IgG) preparations. Two of these preparations contained virus-neutralizing antibodies directed against the IIIb variant's PND. One preparation was a polyclonal chimpanzee antibody purified from an animal previously infected with the IIIb variant and the other preparation was a murine-monoclonal antibody directed to the IIIb

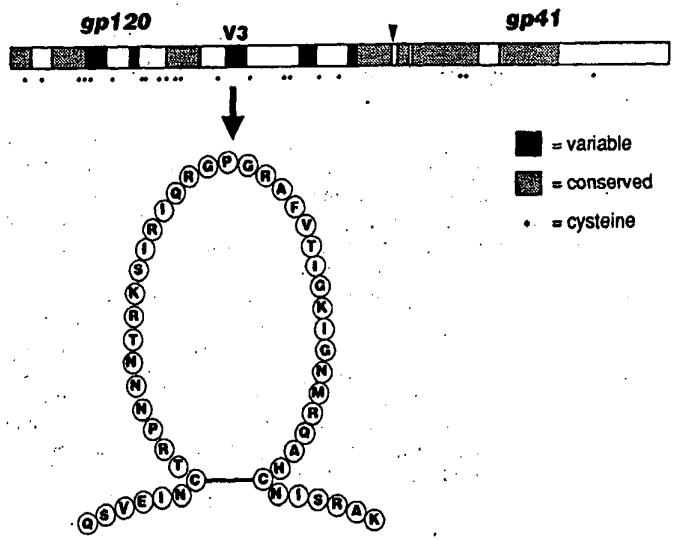


FIGURE 13–1 Schematic representation of the HIV-1 principal neutralization determinant (PND). The determinant encompasses the third hypervariable (V3) domain of the viral gp120 envelope glycoprotein and is known to be a disulfide-bonded, closed loop (see Leonard et al. 1990). The amino acid sequence, in single-letter code, is that of the HIV-1 IIIb variant.

PND. The remaining two IgG preparations were not neutralizing for the IIIb variant of HIV-1, even though one preparation contained antibodies directed against all of the viral structural proteins, but not against the PND. Following a short period of in vitro incubation, each virus-antibody mixture was inoculated into an uninfected chimpanzee. Only the anti-PND virus-neutralizing antibody preparations were found to significantly delay or prevent the establishment of virus infection in the challenged animals.

In addition, Girard et al. (1990) separately immunized two chimpanzees with either inactivated HIV-1 (LAV-1 variant) or with recombinant vaccinia virus expressing the LAV-1 variant gp160 followed by recombinant LAV-1 gp160, core and nonstructural proteins. Neither animal developed virus-neutralizing antibody responses and were therefore subsequently immunized with a synthetic peptide consisting of the LAV-1 variant PND. Measurable neutralizing antibody responses developed and the animals resisted challenge with the closely related IIIb variant virus. Also, Berman et al. (1990) inoculated chimpanzees with a gp120 immunogen that successfully elicited neutralizing antibodies directed against the PND of the IIIb variant

of HIV-1. These chimpanzees were protected following intravenous challenge with this variant of the virus. Finally, two recent studies have shown a correlation between the presence of anti-PND antibodies and the lack of transmission of HIV-1 infection from infected mothers to their infants (Devash et al. 1990; Rossi et al. 1989).

In contrast, chimpanzees immunized with virus envelope glycoprotein-derived or virus core-derived immunogens that did not elicit or elicited very low levels of anti-PND virus-neutralizing antibody at the time of challenge, have not been found to be protected from the establishment of virus infection (Berman et al. 1988; Emini et al. 1990b; Hu et al. 1987; Kennedy et al. 1987; van Eedenburg et al. 1989). In fact, Berman et al. (1990) successfully elicited in chimpanzees a virus-neutralizing antibody response that was not directed against the PND of the challenge virus; Prince et al. (1988) also established an identical immune state in chimpanzees by passive transfer of purified antibody from virus-infected humans. In both cases, the animals were not protected from virus infection following challenge.

Taken together, these results suggest that PND-specific virus-neutralizing antibodies can prevent the in vivo establishment of HIV-1 infection, while neutralizing antibodies directed at non-PND determinants either cannot prevent infection or are much less efficient in doing so. However, further experimentation is required before these conclusions can be firmly stated.

13.4.4 HIV-1 Envelope Glycoprotein-Derived Immunogens

Much of the initial effort directed at the development of an HIV-1 vaccine has concentrated on the virus' external envelope glycoprotein. Robey et al. (1986) purified gp120 from virus-infected cells and, following inoculation into chimpanzees, the immunogen was shown to be a poor inducer of virus-neutralizing antibody (Arthur et al. 1987). Challenge of these animals with live virus, at a time when the elicited neutralizing activity had declined to undetectable levels, demonstrated that a nonneutralizing anti-gp120 response is unprotective (Arthur et al. 1989).

Lasky et al. (1986) reported the expression of gp120 in a genetically engineered Chinese hamster ovary (CHO) cell line. This immunogen, when inoculated into chimpanzees, also proved to be a poor inducer of virus-neutralizing activity (Berman et al. 1988). In a subsequent study, a modification of the inoculation schedule as well as a modification in the preparation of intact gp120 immunogen resulted in better neutralizing antibody titers (Berman et al. 1990). These animals were protected from challenge with the homologous HIV-1 variant and protection appeared to correlate with the level of induced anti-PND antibody (see Section 13.4.3.2). In the same study, chimpanzees inoculated with CHO cell-expressed gp160 envelope glycoprotein precursor also exhibited virus-neutralizing antibody. However, this activity did not appear to be directed against the PND (presumably because the V3 loop had been nonspecifically cleaved during the

glycoprotein's purification) and the animals were not protected upon virus challenge.

gp160 has also been expressed by recombinant vaccinia virus vectors (Dallo et al. 1989; Earl et al. 1989; Hu et al. 1987). These live vectors were capable of eliciting anti-gp160 antibody upon inoculation into chimpanzees, but again the levels of virus-neutralizing antibody were low and the animals were not protected from the establishment of virus infection (Hu et al. 1987; Van Eedenburg et al. 1989). The gp160-expressing vaccinia virus vector has been used in phase I human clinical trials; the initial results of these studies will be discussed later.

gp120 and gp160 have been prepared from recombinant baculovirus expression systems (Rusche et al. 1987; Wells and Compans 1990). One such preparation has been studied in humans in phase I trials and has been shown to elicit several virus-specific immune responses, though the protective efficacy of these responses remains undetermined (Tacket et al. 1990; Orentas et al., 1990).

Finally, various unglycosylated fragments of the envelope glycoprotein have been expressed in recombinant bacteria and yeast expression systems (Barr et al. 1987; Putney et al. 1986). In general, only those fragments containing the PND are capable of inducing significant virus-neutralizing antibodies and, as expected, these antibodies are usually type-specific.

13.4.5 HIV-1 Core Protein-Derived Immunogens

Iwarson et al. (1985) and Murray et al. (1987) independently demonstrated that purified hepatitis B core particles can elicit a protective immune response in chimpanzees against hepatitis B. Hence, an effort was made to determine if the HIV-1 core antigen(s) could also induce such a protective response. Emini et al. (1990b) used yeast-expressed uncleaved HIV-1 core precursor protein, p55, to immunize test chimpanzees. The animals developed an antibody response directed against viral cores, but which was not virus-neutralizing in vitro. The anti-core response did not protect against the establishment of virus infection.

The HIV-1 core proteins have been expressed by recombinant bacteria and by recombinant yeast (Bathurst et al. 1989; Debouck et al. 1987; Graves et al. 1988; Hansen et al. 1988; Jacobs et al. 1989; LeGrice et al. 1988; Vlasuk et al. 1989). The proteins have also been expressed by recombinant baculovirus- and vaccinia-virus-infected cells (Gheysen et al. 1989; Gowda et al. 1989; Flexner et al. 1988; Madisen et al. 1987; Rautmann et al. 1989). In the latter case, the cells elaborate core protein-containing particles (Karacostas et al. 1989; Shioda and Shibuta 1990). Coexpression with the viral envelope glycoprotein leads to the production of particles that contain the processed glycoprotein on their surfaces (Haffar et al. 1990). Immunogenicity studies with these particles have not yet been reported.

13.4.6 Whole Inactivated HIV-1

A number of successful viral vaccines have been developed using inactivated, purified, intact virions as immunogens. Similarly, HIV-1 has been purified and chemically inactivated. The final preparation was composed of intact virions that lacked gp120 on their surfaces. This glycoprotein, being noncovalently attached to gp41, was lost during purification of the virion. As expected, the inactivated immunogen elicited antibody responses in chimpanzees that were directed primarily at the viral core proteins. No virus-neutralizing antibodies were induced and the immune response was not protective against the establishment of virus infection (Gibbs et al. 1989).

13.4.7 PND-specific Immunogens

The chimpanzee studies reported to date suggest that the most important protective immune response against HIV-1 is antibody directed to the gp120 PND. However, the PND has proven to be poorly immunogenic in the context of the intact glycoprotein or gp120 fragments. Also, the variability of the V3 domain's amino acid sequence has limited the antibody response to this site in a variant-specific fashion. These immunogenicity and variability concerns have made the PND a prime candidate for the development of synthetic immunogens, which, if appropriately constructed, can address both problems.

Synthetic peptides containing PND sequences have been used successfully to elicit virus-neutralizing antibody responses, but usually these responses are type-specific. In an attempt to overcome this specificity and define the extent of PND variability, LaRosa et al. (1990) determined and compared the sequences of 245 PNDs from primary HIV-1 isolates and laboratory variants of the virus. Most of the PND's variability appears to be clustered in two regions on either side of the amino acid residues at the "tip" of the PND (Figure 13–2). The tip exhibits a certain degree of conservation. The sequence I-G-P-G-R-A-F is found in over 50% of the sequences and the majority of the sequences are similar to a laboratory variant of HIV-1 called MN (Gurgo et al. 1988). Sequences similar to the prototypic IIIb and LAV-1_Bru variants, which were used for most of the biological studies performed to date, represent less than 5% of the total. This sequence distribution agrees with the results of serological studies which showed that 65% of HIV-seropositive humans have antibodies that bind synthetic peptide representations of the MN variant PND, whereas less than 15% have antibodies that bind the IIIb variant PND (LaRosa et al. 1990). In addition, the PNDs of the HIV-1 variants appear to exhibit certain common structural features, particularly around the loop tip region.

These somewhat conserved structural and sequence features establish the hope that synthetic peptide immunogens may be developed that will preferentially drive antibody responses to the conserved elements and,

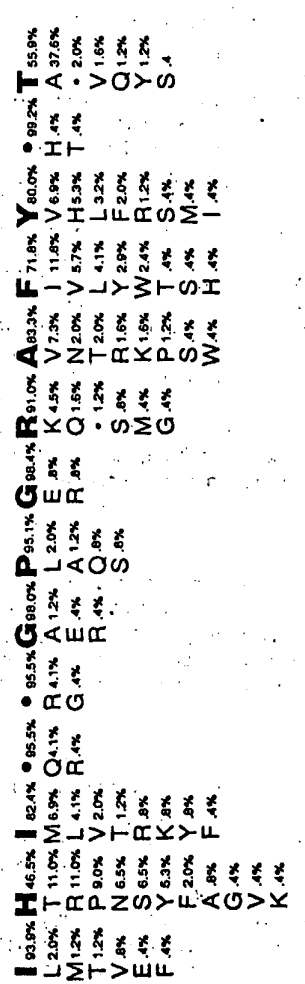


FIGURE 13–2 Amino acid sequence variability in the central region of the PND. The amino acids found at each residue position are listed along with the frequency with which each occurs. The top sequence (bold letters) represents a consensus sequence composed of the most commonly occurring amino acid at each position. The data for this figure were obtained from the PND sequences of 245 HIV-1 variants (LaRosa et al. 1990). The single-letter amino acid code is used.

hence, elicit broadly reactive neutralizing activity. Such synthetic peptides, if constructed with the appropriate carriers and formulated with the appropriate vehicles, may also serve to elicit neutralizing antibody titers of acceptable levels and significant duration. A number of efforts currently are in progress to develop and test such PND-specific immunogens.

An example of this approach is that described by Palker et al. (1989) in which a PND peptide was covalently synthesized onto a peptide containing an identified T-helper sequence from the gp120 glycoprotein. Two T-helper determinants in the gp120 have been identified by Berzofsky et al. (1988). A synthetic immunogen containing the PND and a gp120-specific T-helper determinant would not only elicit anti-PND antibody, but may prime the immune system for an anamnestic antibody response following the initial virus infection. If the persistent viral infection were to be established only after an initial period of virus replication, then such an anamnestic response might be protective.

13.5 HIV-1 VACCINES: CLINICAL STUDIES

Several phase I human clinical studies with HIV-1 vaccine candidates are in progress to assess safety and immunogenicity. Given the absence of an adequate understanding of the immunological basis of HIV-1 vaccine efficacy, emphasis was placed on eliciting immune responses directed to the viral envelope glycoprotein.

gp160 expressed in a recombinant baculovirus system has been found to be immunogenic at doses greater than 100 μg for the elicitation of antienvelope glycoprotein antibodies and for the induction of virus-neutralizing activity (Smith and Volvovitz 1990). The immunogen, in a fraction of the volunteer recipients, also primed for an MHC class II-restricted CTL response against virus-infected cells mediated by CD4-positive cytotoxic T-cells (Orentas et al. 1990). The protective efficacy of this immune response and of the antibody responses to the gp120 immunogen remain unclear. Protective efficacy studies with this immunogen in chimpanzees have not been reported.

Live vaccinia virus vectors expressing the HIV-1 gp160 also have been tested in phase I studies (Cooney et al. 1990). The recombinant virus proved to be moderately immunogenic for the induction of anti-gp160 antibody responses when inoculated into vaccinia-naive recipients. The recombinant virus was not immunogenic for anti-gp160 responses if inoculated into individuals who had previously been vaccinated with the vaccinia virus. Similarly, the recombinant virus was not effective if inoculated a second time to boost the anti-gp160 antibody titers. Presumably the rapidly developing antivaccinia responses prevent sufficient replication of the recombinant virus to allow for the required expression of gp160. An effort to overcome this restriction has been made in which volunteer recipients were first inoculated with the vaccinia recombinant and then were boosted with the baculovirus-expressed, purified HIV-1 envelope glycoprotein. This resulted in somewhat better anti-gp160 antibody responses, though the inducton of virus-neutralizing activity remained poor.

Finally, clinical studies are in progress with recombinant yeast-derived nonglycosylated fragments of gp120 using a muramyl dipeptide adjuvant. Complete results from these studies have not yet been reported.

13.6 BASIC VACCINE STUDIES WITH THE SIMIAN IMMUNODEFICIENCY VIRUS

The simian immunodeficiency viruses (SIVs) are related to HIV-1 and have been isolated from a number of diverse monkey species. Certain variants of this group of viruses, when inoculated into the appropriate primate species, will mediate the development of an AIDS-like clinical disease after a period of persistent infection. Given the logistical difficulties of the chimpanzee model of HIV-1 infection, many researchers have focused on the SIV monkey model to perform vaccine studies. It is too early to judge whether the results of these studies will provide information that will be applicable to the development of an HIV-1 vaccine since the SIVs and HIV-1 are immunologically distinct. The SIVs, for example, do not have a readily identifiable PND that is immunologically and structurally homologous to the HIV-1 PND. However, attempts to establish an immune state in monkeys capable of preventing an SIV infection have been successful and, therefore, provide additional support for believing that this goal may be practically attainable in the HIV-1 system.

Desrosiers et al. (1989) and Murphy-Corb et al. (1989) used chemically inactivated, whole SIV to immunize rhesus monkeys that were then challenged with live virus. In both studies, the immunization clearly resulted in protection against the establishment of the persistent virus infection when low-challenge doses of the virus were used. The immunological basis for the protection is not known. The protected animals exhibited little or no in vitro virus-neutralizing activity at the time of challenge.

Also, in these studies and in those reported by Sutjipto et al. (1990), a proportion of the monkeys was not protected from infection following immunization. However, a delay was seen in the development of the immunodeficient state. Analysis of the quantity of virus present in the lymph node tissues after infection showed that the immunized animals contained notably lower levels of the virus than did the nonimmunized controls (Desrosiers et al. 1989). A similar delay in disease progression was reported by Marthas et al. (1990) following virulent virus challenge of monkeys that had been immunized with a live, attenuated SIV variant. Again, the immunological basis for this apparent partial-protection is unknown.

13.7 SUMMARY AND CONCLUSIONS

A successful AIDS vaccine must elicit an immune state that will prevent the establishment of an HIV-1 persistent infection. This is a unique and difficult goal for a vaccine. Most vaccines elicit or prime for immune responses that prevent or attenuate the expression of clinical disease following infection with the pathogen. However, current evidence suggests that, following persistent infection with HIV-1, antiviral immune responses do not prevent the long-term progression to disease. Hence, it seems that the development of the persistent infection must be prevented.

The ability of the immune response to accomplish this goal depends upon the efficiency with which the virus establishes persistence in the host. This is unknown for HIV-1. As a result, early efforts at vaccine development have focused on humoral immune responses directed against the virus particle in the attempt to prevent any infection of the host's cells. Studies with chimpanzees, as a model for HIV-1 infection, suggest that virus-neutralizing antibodies directed against the third hypervariable (V3) domain of the viral gp120 envelope glycoprotein may be particularly effective in preventing this infection. Studies also are in progress, both in chimpanzees and humans, to define the immunogenicity and effectiveness of various immunogens derived from the viral envelope and core structural proteins.

Efforts that have concentrated on the gp120 V3 domain (or PND) have defined the extent of this region's variability and have established elements of generally conserved structure and sequence. The construction of these elements into practical and effective immunogens is an important goal.

Finally, it is essential that basic studies be performed to determine if humoral or cellular immune responses directed against virus-infected cells would aid in preventing the establishment of an HIV-1 persistent infection. Such immune responses, if effective and in conjunction with specific virus-neutralizing antibody responses, would enhance the probability that an effective HIV-1 vaccine could be developed.

REFERENCES

Allan, J.S., Coligan, J.E., Barin, F., et al. (1985) *Science* 228, 1091–1094.
Alter, H.J., Eichberg, J.W., Masur, H., et al. (1984) *Science* 226, 549–552.
Arthur, L.O., Bess, J.W., Waters, D.J., et al. (1989) *J. Virol.* 63, 5046–5053.
Arthur, L.O., Pyle, S.W., Nara, P.L., et al. (1987) *Proc. Natl. Acad. Sci. USA* 84, 8583–8587.
Barr, P.J., Steimer, K.S., Sabin, E.A., et al. (1987) *Vaccine* 5, 90–101.
Bathurst, I.C., Chester, N., Gibson, H.L., et al. (1989) *J. Virol.* 63, 3176–3179.
Berkelman, R.L., and Curran, J.W. (1989) *Epidemiol. Rev.* 11, 222–228.
Berkower, I., Smith, G.E., Giri, C., and Murphy, D. (1989) *J. Exp. Med.* 170, 1681–1695.
Berman, P.W., Gregory, T.J., Riddle, L., et al. (1990) *Nature* 345, 622–625.

Berman, P.W., Groopman, J.E., Gregory, T., et al. (1988) *Proc. Natl. Acad. Sci. USA* 85, 5200–5204.
Berzofsky, J.A., Bensussan, A., Cease, L.B., et al. (1988) *Nature* 334, 706–708.
Centers for Disease Control (1990a) *Morbidity Mortality Weekly Rep.* 39, 81–86.
Centers for Disease Control (1990b) *Morbidity Mortality Weekly Rep.* 39, 110–119.
Chanh, T.C., Dreesman, G.R., Kanda, P., et al. (1986) *EMBO J.* 5, 3065–3071.
Clavel, F., Guetard, D., Brun-Vezinet, F., et al. (1986) *Science* 233, 343–346.
Cooney, E.L., Corey, L., Hu, S.L., et al. (1990) in *6th International Conference on AIDS*, abstr. Th.A.333.
Dalgleish, A.G., Beverly, P.C.L., Clapham, P.R., et al. (1984) *Nature* 312, 763–768.
Dalgleish, A.G., Chanh, T.C., Kennedy, R.C., et al. (1988) *Virol.* 165, 209–215.
Dallo, S., Maa, J.S., Rodriguez, J.R., et al. (1989) *Virol.* 173, 323–329.
Daniel, M.D., Letvin, N.L., King, N.W., et al. (1985) *Science* 228, 1201–1204.
Debouck, C., Gorniak, J.G., Strickler, J.E., et al. (1987) *Proc. Natl. Acad. Sci. USA* 84, 8903–8906.
Desrosiers, R.C., Wyand, M.S., Kodama, T., et al. (1989) *Proc. Natl. Acad. Sci. USA* 86, 6353–6357.
Devash, Y., Calvelli, T.A., Wood, D.G., Reagan, K.J., and Rubinstein, A. (1990) *Proc. Natl. Acad. Sci. USA* 87, 3445–3449.
Earl, P.L., Robert-Guroff, M., Matthews, T.J., et al. (1989) *AIDS Res. Hum. Retroviruses* 5, 23–32.
Emini, E.A., Nara, P.L., Schleif, W.A., et al. (1990a) *J. Virol.* 64, 3674–3678.
Emini, E.A., Schleif, W.A., Quintero, J.C., et al. (1990b) *AIDS Res. Hum. Retroviruses* 6, 1247–1250.
Evans, D.J., McKeating, J., Meredith, J.M., et al. (1989) *Nature* 339, 385–388.
Fauci, A.S. (1988) *Science* 239, 617–622.
Filice, G., Cereda, P.M., and Varnier, O.E. (1988) *Nature* 335, 366–369.
Flexner, C., Brogles, S.S., Ear, P., Chakrabarti, S., and Moss, B. (1988) *Virol.* 166, 339–349.
Fung, M.S.C., Sun, C., Sun, N.C., Chang, N.T., and Chang, T.W. (1987) *Biotechnol.* 5, 940–946.
Gheysen, D., Jacobs, E., deForesta, F., et al. (1989) *Cell* 59, 103–112.
Gibbs, C.J., Mora, C., Peters, R., et al. (1989) in *5th International AIDS Conference*, abstr. Th.C.0.46.
Girard, M., Kaczorek, M., Pinter, A., et al. (1990) *J. Cell. Biochem.* (suppl. 14D), 150.
Gonda, M.A., Braun, M.J., Carter, S.G., et al. (1988) *Nature* 330, 388–391.
Goudsmit, J., Boucher, C.A.B., Meloen, R.H., et al. (1988a) *AIDS* 2, 157–164.
Goudsmit, J., Debouck, C., Meloen, R.H., et al. (1988b) *Proc. Natl. Acad. Sci. USA* 85, 4478–4482.
Goudsmit, J., Thiriart, C., Smit, C., Bruck, C., and Gibbs, C.J. (1988c) *Vaccine* 6, 229–232.
Gowda, S.D., Stein, B.S., Steimer, K.S., and Engleman, E.G. (1989) *J. Virol.* 63, 1451–1454.
Graves, M.C., Lim, J.T., Heimer, J.P., and Kramer, R.A. (1988) *Proc. Natl. Acad. Sci. USA* 85, 2449–2453.
Gurgo, C., Guo, H.G., Franchini, G., et al. (1988) *Virol.* 164, 531–536.
Haffar, O., Garrigues, J., Travis, B., et al. (1990) *J. Virol.* 64, 2653–2659.
Hahn, B.H., Shaw, G.M., Taylor, M.E., et al. (1986) *Science* 232, 1548–1553.

Haigwood, N.L., Barker, C.B., Higgins, K.W., et al. (1990) in *Vaccines 90* (Chanock, R.M., Lerner, R.A., Brown, F., and Ginsberg, H., eds.), pp. 313–320, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Hansen, J., Bilich, S., Schulze, T., Sukrow, S., and Moelling, K. (1988) *EMBO J.* 7, 1785–1791.
Ho, D.D., Kaplan, J.C., Rackauskus, I.E., and Gurney, M.E. (1988) *Science* 239, 1021–1023.
Ho, D.D., Sarngadharan, M.G., Hirsch, M.S., et al. (1987) *J. Virol.* 61, 2024–2028.
Hosmalin, A., Clerici, M., Houghten, R., et al. (1990) *Proc. Natl. Acad. Sci. USA* 87, 2344–2348.
Hu, S.L., Fultz, P.N., McClure, H.M., et al. (1987) *Nature* 328, 721–723.
Iwarson, S., Tabor, E., Thomas, H.C., Snog, P., and Gerety, R.J. (1985) *Gastroenterol.* 88, 763–767.
Jacobs, E., Gheysen, D., Thines, D., Francotte, M., and deWilde, M. (1989) *Gene* 79, 71–78.
Javaherian, K., Langlois, A.J., McDanal, C., et al. (1989) *Proc. Natl. Acad. Sci. USA* 86, 6768–6772.
Karacostas, V., Nagashima, K., Gouda, M.A., and Moss, B. (1989) *Proc. Natl. Acad. Sci. USA* 86, 8964–8967.
Kenealy, W.R., Matthews, T.J., Ganfield, M.C., et al. (1989) *AIDS Res. Hum. Retroviruses* 5, 173–182.
Kennedy, R.C., Kanda, P., Dreesman, G.R., Eichberg, J.W., and Chanh, T.C. (1987) *Vaccines 87* (Chanock, R.M., Lerner, R.A., Brown, F., and Ginsberg, H., eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Klatzmann, D., Champagne, E., Chamaret, S., et al. (1984) *Nature* 312, 767–768.
Koenig, S., Earl, P., Powell, D., et al. (1988) *Proc. Natl. Acad. Sci. USA* 85, 8638–8642.
Kohl, N.E., Emini, E.A., Schleif, W.A., et al. (1988) *Proc. Natl. Acad. Sci. USA* 85, 4686–4690.
Kowalski, M., Potz, J., Basiripour, L., et al. (1987) *Science* 237, 1351–1355.
Kulaga, H., Folks, T.M., Rutledge, R., and Kindt, T.J. (1988) *Proc. Natl. Acad. Sci. USA* 85, 4455–4459.
Kulaga, H., Folks, T., Rutledge, R., et al. (1989) *J. Exp. Med.* 162, 321–326.
LaRosa, G.J., Davide, J.P., Weinhold, K., et al. (1990) *Science* 249, 932–935.
Lasky, L.A., Groopman, J.E., Fennie, C.W., et al. (1986) *Science* 233, 209–212.
LeGrice, S.F.J., Mills, J., and Moss, J. (1988) *EMBO J.* 7, 2547–2553.
Leonard, C.K., Spellman, M.W., Riddle, L., et al. (1990) *J. Biol. Chem.* 265, 10373–10382.
Linsley, P.S., Ledbetter, J.A., Thomas, E.K., and Hu, S.L. (1988) *J. Virol.* 62, 3695–3702.
Ljunggren, K., Bottiger, B., Biberfeld, B., et al. (1987) *J. Immunol.* 139, 2263–2267.
Madisen, L., Travis, B., Hu, S.L., and Purchio, A.F. (1987) *Virol.* 158, 248–250.
Marthas, M.L., Sutjipto, S., Higgins, J., et al. (1990) *J. Virol.* 64, 3694–3700.
Matsushita, S., Robert-Guroff, M., Rusche, J., et al. (1988) *J. Virol.* 62, 2107–2114.
Matthews, T.J., Langlois, A.L., Robey, W.G., et al. (1986) *Proc. Natl. Acad. Sci. USA* 83, 9709–9713.
McChesney, M., Tanneau, F., Regnault, A., et al. (1990) *Eur. J. Immunol.* 20, 215–220.
Modrow, S., Hahn, B.H., Shaw, G.M., et al. (1987) *J. Virol.* 61, 570–578.

Mosier, D.E., Gulizia, R.J., Baird, S.M., and Wilson, D.B. (1988) *Nature* 335, 256–259.
Muesing, M.A., Smith, D.H., Cabradilla, C.D., et al. (1985) *Nature* 313, 450–458.
Murphy-Corb, M., Martin, L.N., Davison-Fairburn, B., et al. (1989) *Science* 246, 1293–1297.
Murphy-Corb, M., Martin, L.N., Rangan, S.R.S., et al. (1986) *Nature* 321, 435–437.
Murray, K., Bruce, S.A., Wingfield, P., et al. (1987) *J. Med. Virol.* 23, 101–107.
Namikawa, R., Kaneshima, H., Lieberman, M., Weissman, I.L., and McCune, J.M. (1988) *Science* 242, 1684–1686.
Nara, P.L. (1989) in *Vaccines 89* (Chanock, R.M., Lerner, R.A., Brown, F., and Ginsberg, H., eds.), pp. 137–144, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Nara, P.L., Robey, W.G., Arthur, L.O., et al. (1987) *J. Virol.* 61, 3173–3180.
Nara, P.L., Robey, W.G., Pyle, S.W., et al. (1988) *J. Virol.* 62, 2622–2628.
Nixon, D.F., Townsend, A.R., Elvin, J.G., et al. (1988) *Nature* 336, 484–487.
Orentas, R.J., Hildreth, J.E.K., Obah, E., et al. (1990) *Science* 248, 1234–1237.
Palker, T.J., Clark, M.E., Langlois, A.J., et al. (1988) *Proc. Natl. Acad. Sci. USA* 85, 1932–1936.
Palker, T.J., Matthews, T.J., Langlois, A., et al. (1989) *J. Immunol.* 142, 3612–3619.
Pavlakis, G.N., and Felber, B.K. (1990) *New Biol.* 2, 20–31.
Payne, S.L., Salinovich, O., Montelaro, R.C., Issel, C.J., and Nauman, S.M. (1987) *J. Virol.* 61, 1266–1270.
Pedersen, N.C., Ho, E., Brown, M.L., and Yamamoto, J.K. (1987) *Science* 235, 790–793.
Plata, F., Autran, B., Martins, L.P., et al. (1987) *Nature* 328, 348–351.
Prince, A.M., Horowitz, B., Baker, L., et al. (1988) *Proc. Natl. Acad. Sci. USA* 85, 6944–6948.
Putney, S.D., Matthews, T.J., Robey, W.G., et al. (1986) *Science* 234, 1392–1395.
Ratner, L., Haseltine, W., Patarca, R., et al. (1985) *Nature* 313, 277–284.
Rautmann, G., Kieny, M.P., Brandely, R., et al. (1989) *AIDS Res. Hum. Retroviruses* 5, 147–157.
Ricchetti, M., and Buc, H. (1990) *EMBO J.* 9, 1583–1593.
Riviere, Y., Tanneau-Salvadori, F., Regnault, A., et al. (1989) *J. Virol.* 63, 2270–2277.
Robey, W.G., Arthur, L.O., Matthews, T.J., et al. (1986) *Proc. Natl. Acad. Sci. USA* 83, 7023–7027.
Rossi, P., Moschese, V., Broliden, P.A., et al. (1989) *Proc. Natl. Acad. Sci. USA* 86, 8055–8058.
Rusche, J.R., Javaherian, K., McDanal, C., et al. (1988) *Proc. Natl. Acad. Sci. USA* 85, 3198–3202.
Rusche, J.R., Lynn, D.L., Robert-Guroff, M., et al. (1987) *Proc. Natl. Acad. Sci. USA* 84, 6924–6928.
Shepp, D.H., Daguillard, F., Mann, D., and Quinnan, G.V. (1988) *AIDS* 2, 113–117.
Shioda, T., and Shibuta, H. (1990) *Virol.* 175, 139–148.
Skinner, M.A., Langlois, A.J., McDanal, C.B., et al. (1988a) *J. Virol.* 62, 4195–4200.
Skinner, M.A. Ting, R., Langlois, A.J., et al. (1988b) *AIDS Res. Hum. Retroviruses* 4, 187–197.
Smith, G., and Volvovitz, F. (1990) in *6th International Conference on AIDS*, abstr. S.A.75.

Sutjipto, S., Pedersen, N.C., Gardner, M.B., et al. (1990) *J. Virol.* 64, 2290–2297.
Tacket, C.O., Baqar, S., Munoz, C., and Murphy, J.R. (1990) *AIDS Res. Hum. Retroviruses* 6, 535–542.
Takahashi, H., Cohen, J., Hosmalin, A., et al. (1988) *Proc. Natl. Acad. Sci. USA* 85, 3105–3109.
Thomas, S.K., Weber, J.N., McClure, J., et al. (1988) *AIDS* 2, 25–29.
Van Eedenburg, J.P., Yagello, M., Girard, M., et al. (1989) *AIDS Res. Hum. Retroviruses* 5, 41–50.
Vlasuk, G.P., Waxman, L., Davis, L.J., et al. (1989) *J. Biol. Chem.* 264, 12106–12112.
Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S., and Alizon, M. (1985) *Cell* 40, 9–17.
Walker, B.D., Chakrabarti, S., Moss, B., et al. (1987) *Nature* 328, 345–348.
Walker, B.D., Flexner, C., Birch-Limberger, K., et al. (1989) *Proc. Natl. Acad. Sci. USA* 86, 9514–9518.
Walker, B.D., Flexner, C., Paradis, T.J., et al. (1988) *Science* 240, 64–66.
Wells, D.E., and Compans, R.W. (1990) *Virol.* 176, 575–586.
Zagury, D., Bernard, J., Cheynier, R., et al. (1988) *Nature* 332, 728–731.

CHAPTER
14

Recent Advances in Antitumor Vaccines

Shiu-Lok Hu
Ingegerd Hellström
Karl Erik Hellström

14.1 IMMUNOLOGICAL BASIS FOR ANTITUMOR VACCINES

Many experimentally induced animal tumors express tumor-specific transplantation antigens (TSTAs) that can act as targets of an immune response that leads to the rejection of neoplastic cells transplanted onto properly immunized syngeneic hosts. Since immunization against TSTAs can sometimes cause the destruction also of small, already established tumors, TSTAs might be used as therapeutic immunogens (vaccines) to induce a tumoricidal immune response.

The immune response causing tumor destruction is primarily mediated by lymphocytes (Hellström, K.E., and Hellström 1969), most notably T-cells (Greenberg 1991). As such, immune memory can be activated to kill tumor cells that may arise a long time after the therapy has been completed.

We wish to thank Dr. Jim Lewis for critical reading of the manuscript, and Mrs. Phyllis Harps and Ms. Ana Wieman for manuscript preparations.

JOURNAL OF VIROLOGY, Nov. 1993, p. 6897–6902
0022-538X/93/116897-06$02.00/0

Vol. 67, No. 11

Mutations in Human Immunodeficiency Virus Type 1 gp41 Affect Sensitivity to Neutralization by gp120 Antibodies

NICOLE K. T. BACK,[1]* LIA SMIT,[1] MARTIN SCHUTTEN,[2] PETER L. NARA,[3] MATTHIJS TERSMETTE,[1] AND JAAP GOUDSMIT[1]

Department of Virology, Academic Medical Center, 1105 AZ Amsterdam,[1] and Laboratory of Immunobiology, National Institute of Public Health and Environmental Protection, 3720 BA Bilthoven,[2] The Netherlands, and Laboratory of Tumor Cell Biology, National Cancer Institute-Frederick Cancer Research and Development Center, Frederick, Maryland 21701-1013[3]

Received 17 December 1992/Accepted 10 August 1993

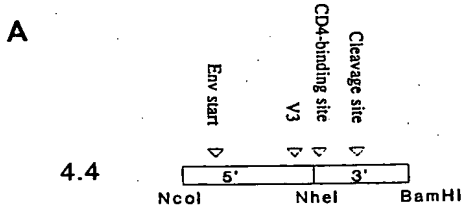
Three closely related molecular human immunodeficiency virus type 1 (HIV-1) clones, with differential neutralization phenotypes, were generated by cloning of an *Nco*I-*Bam*HI envelope (*env*) gene fragment (HXB2R nucleotide positions 5221 to 8021) into the full-length HXB2 molecular clone of HIV-1 IIIB. These *env* gene fragments, containing the complete gp120 coding region and a major part of gp41, were obtained from three different biological clones derived from a chimpanzee-passaged HIV-1 IIIB isolate. Two of the viruses thus obtained (4.4 and 5.1) were strongly resistant to neutralization by infection-induced chimpanzee and human polyclonal antibodies and by HIV-1 IIIB V3-specific monoclonal antibodies and weakly resistant to soluble CD4 and a CD4-binding-site-specific monoclonal antibody. The third virus (6.8) was sensitive to neutralization by the same reagents. The V3 coding sequence and the gp120 amino acid residues important for the discontinuous neutralization epitope overlapping the CD4-binding site were completely conserved among the clones. However, the neutralization-resistant clones 4.4 and 5.1 differed from neutralization-sensitive clone 6.8 by two mutations in gp41. Exchange experiments confirmed that the 3′ end of clone 6.8 (nucleotides 6806 to 8021; amino acids 346 to 752) conferred a neutralization-sensitive phenotype to both of the neutralization-resistant clones 4.4 and 5.1. From our study, we conclude that mutations in the extracellular portion of gp41 may affect neutralization sensitivity to gp120 antibodies.

Experimental human immunodeficiency virus type 1 (HIV-1) infection of chimpanzees and natural HIV-1 infection of humans induce type-specific as well as group-specific neutralizing antibodies in the course of infection (3, 9, 10, 18, 25, 33). Viruses resistant to neutralization by autologous sera emerge during the course of human HIV-1 infection (1, 16, 29, 31, 32). HIV-1 variants with reduced sensitivity to neutralization by autologous but not by heterologous broadly neutralizing sera occur as early as 1 year after infection (1, 31). Despite the V3-dependent restriction of neutralizing capacity of sera during the early phase of infection, the V3 sequence of the neutralization-resistant viruses appeared to be unchanged (32). These findings in humans confirmed our earlier report on the emergence of neutralization-resistant virus variants in chimpanzees experimentally infected with HIV-1 IIIB (19). Virus recovered from a chimpanzee and passaged 32 weeks after infection had a reduced sensitivity to V3-specific experimental as well as autologous chimpanzee sera, although the amino acid sequence of the neutralizing epitope within the V3 loop was unchanged (19). Reitz and coworkers have shown that human neutralizing serum can select in vitro a mutant resistant to the selecting serum as a result of a mutation in gp41 (24, 35). Although the specificity of the antibodies in the serum driving this selection is not known, this gp41 mutation appears to indirectly affect the neutralization epitopes of the CD4-binding site of gp120 or epitopes overlapping with this region (26b). In this study, we further analyzed the genetic background of changes in neutralization phenotype due to mutations outside the primary antibody-binding site and the possible mechanism responsible for these changes in neutralization phenotype. For this purpose, we studied three molecular clones differing in neutralization sensitivity, obtained by cloning of *env* gene fragments derived from a chimpanzee-passaged HIV-1 IIIB isolate (911.32) (19).

Construction of envelope molecular clones. A chimpanzee-passaged virus variant, obtained 32 weeks after infection with HIV-1 IIIB, was isolated from chimpanzee 911 peripheral blood mononuclear cells (19). This bulk virus isolate showed reduced sensitivity to autologous sera and V3 monoclonal antibody (MAb) 0.5β, while an identical V3 region (λBH10) was found (19). *env* gene fragments of biological clones of the bulk isolate, obtained by two rounds of plaque purification on MT-4 cells and propagated in the C8166 cell line (15), were cloned into a full-length HXB2 molecular clone (6). For this cloning procedure, DNA of the biologically cloned viruses was extracted (4) and amplified by the polymerase chain reaction. Primers used to amplify the envelope region were an extended 5′ primer, pol (A*) (5′-AAGCTNCTCTGGAAAGGTGAAG GGGCAGTA-3′; HXB2R nucleotides [nt] 4493 to 4522) (17), and a 3′ primer, nef.seq.3 (5′-CTACTTGTGATTGCTCC ATG-3′; HXB2R nt 8460 to 8479) (17). The reaction mixture (100 μl) consisted of 25 mM KCl, 50 mM Tris-HCl (pH 8.3), 0.1 mg of bovine serum albumin per ml, 2 mM $MgCl_2$, 0.2 mM (each) deoxynucleoside triphosphate, 10 pmol of each primer, 5% glycerol, and 2.5 U of *Taq* polymerase (a gift from Perkin-Elmer Cetus, Emeryville, Calif.). Each cycle consisted of 1 min at 94°C, 2 min at 55°C, and 4 min at 72°C. After 35 cycles, the reactions were extended for 10 min at 72°C. The amplified envelope fragments were purified and digested with *Nco*I (nt 5221) and *Bam*HI (nt 8021). The *Nco*I-*Bam*HI fragment was subsequently cloned in a full-length HXB2 molecular clone lacking the *Nco*I-*Bam*HI fragment (6). Of the three infectious molecular clones thus obtained, four chimeric

* Corresponding author.

A



B

	Chimpanzee		Human	CD4-site		V3-site	
	36wk	87wk	1171	GP13	sCD4	178.1	5023
4.4	20	10	<10	<5	25	20	40
5.1	<20	10	<10	40	50	80	80
6.8	160	80	320	400	100	1600	800
4.4 - 6.8	640	ND	640	800	200	1600	800
5.1 - 6.8	450	ND	640	650	100	6400	3200
6.8 - 4.4	<20	ND	<10	280	50	100	200
6.8 - 5.1	40	ND	10	570	50	140	200

FIG. 1. (A) Schematic representation of the three molecular clones with different envelope regions and two-thirds of the gp41 protein and of the four exchanged chimeric molecular clones. A 4-kb fragment was obtained by the polymerase chain reaction, digested with *Nco*I and *Bam*HI, and cloned in a full-length HXB2 molecular clone. Exchange experiments were performed with *Nhe*I and *Bam*HI fragments. Depicted are the most important functional regions (*env* start region, neutralization V3 epitope, CD4-binding region, and gp120-gp41 cleavage site) and the restriction enzyme sites used (*Nco*I, *Nhe*I, and *Bam*HI). (B) Reciprocal neutralizing titers of the chimeric molecularly cloned viruses against four sets of sera. The first two sets are polyclonal antibodies; the other two sets are MAbs.

envelope molecular clones were prepared by exchange of the *Nhe*I-*Bam*HI fragment (nt 6806 to 8021; amino acids [aa] 346 to 752) (Fig. 1).

Viral stock preparation for neutralization experiments. After cesium chloride gradient purification of the full-length molecular clone DNA, 5 μg of each molecular clone was electroporated into 5×10^6 SupT1 cells and maintained as described by de Jong et al. (6). On day 3, an equal amount of fresh medium, RPMI 1640 (GIBCO, Paisley, Scotland) supplemented with 10% fetal calf serum, glutamine, and antibiotics, was added. On day 5, the supernatant fluids of the transfected cells were harvested by centrifugation at $1,000 \times g$ for 10 min and filtered through a Millipore filter (0.22-μm pore size). Supernatant fluids of HIV-1 IIIB variants and 911.32 chronically infected stocks were harvested as described previously (2). Infectious titers were determined by endpoint titration on the basis of syncytium formation, using the C8166 cell line as previously described (2, 15).

Virus neutralization assay. The cell-free virus neutralization assay was performed as described previously (2, 15). Briefly, 100 50% tissue culture infectious doses of virus stock in a volume of 40 μl was incubated in triplicate with 10 μl of twofold serial dilutions of antibodies and incubated at 37°C for 1 h. The virus-antibody mixture was then incubated with 100 μl of C8166 cells (2×10^5 cells per ml) and scored for syncytia at day 5. Neutralization titer is represented as the reciprocal of the highest serum dilution giving 99% inhibition of syncytium formation.

Sera and monoclonal antibodies. The following polyclonal and monoclonal antibodies were used to characterize the molecularly cloned viruses. Autologous plasma samples of chimpanzee 911, obtained 36 and 87 weeks after infection with HIV-1 IIIB, contain both antibodies to V3 of HIV-1 HX10 and antibodies able to block gp120-soluble CD4 (sCD4) binding and can neutralize HIV-1 HX10 and RF but not MN (data not shown); a human serum from the Dutch cohort (23), selected for high neutralizing titers against the laboratory strains HIV-1 MN, IIIB, and RF, contains HIV-1 MN V3 but not HIV-1 HX10 V3 antibodies and also contains gp120-GP13 blocking antibodies (data not shown). Human neutralizing MAb GP13, derived by Epstein-Barr virus transformation of peripheral blood mononuclear cells from an asymptomatic, HIV-1-seropositive individual, was used as a protein A-purified preparation (1 mg/ml). This human MAb binds to an epitope overlapping that of the CD4-binding site (26a). Recombinant sCD4 (1 mg/ml) was kindly provided by P. J. Maddon through the NIAID AIDS Research and Reference Reagent Program. We also used two HIV-1 IIIB V3-specific mouse MAbs, 5023 and 178.1. MAb 5023 was induced by a 15-mer synthetic BH10 V3 peptide (7), specific for RgPGRAF (aa 311 to 317; capital letters represent residues important in binding) (13), kindly provided both as ascites fluid and as a batch of purified immunoglobulin G (1 mg/ml) by P. Durda (Du Pont de Nemours and Co., North Billerica, Mass.). The recombinant gp160 (IIIB) DNA-induced MAb 178.1 (28), specific for KSiRI (aa 305 to 309) (14), was kindly provided as ascites fluid by C. Thiriart (Smith Kline, Rixensart, Belgium) and by the MRC AIDS Reagent Project.

TABLE 1. Reciprocal neutralizing titers of three uncloned HIV-1 IIIB virus isolates and the chimpanzee-passaged virus bulk isolate (911.32) against two V3-specific MAbs and a human broadly neutralizing serum

HIV-1 IIIB isolate	Titer[a]		
	V3 site MAb		1171 (human serum)
	178.1[b]	5023[b]	
HX10	520	6,400	80
HXB2	<50	570	28
HXB3	<10	1,131	10
911.32 bulk	<50	<50	<10

[a] Logarithmic mean of two or more separate experiments.
[b] Ascites fluid.

Sequencing of the *env* coding regions of the three molecular clones. Nucleotide sequence analysis of the envelope genes was performed by the Central European Sequence Facility for HIV genome analysis (H. Rübsamen, Georg-Speyer Haus, Frankfurt, Germany) by an automated sequence method. All nucleotide sequence mutations were confirmed by an automated sequence method (Applied Biosystems, Maarssen, The Netherlands), using the *Taq* dye terminator cycle kit (Applied Biosystems) and various envelope primers synthesized by the Central Laboratory of the Netherlands Red Cross Blood Transfusion Service, Amsterdam, The Netherlands.

Neutralization profiles of chimpanzee-passaged virus versus HIV-1 IIIB molecular clones. Polyclonal and HIV-1 IIIB V3-specific MAbs were tested to determine the neutralization sensitivity of the three HIV-1 IIIB variants (HX10, HXB2, and HXB3) and the chimpanzee passage of HIV-1 IIIB (911.32). As shown in Table 1, the V3-binding MAbs were able to neutralize HX10 but were substantially less able to neutralize 911.32. As previously shown for the V3-specific MAb 0.5β, isolate 911.32 was also resistant to neutralization by two other V3 MAbs, 178.1 and 5023, used in this study (Table 1). This neutralization resistance could not be attributed to changes in the binding sites of these V3 MAbs (13, 14, 19), as in the core for clones HXB2 and HXB3, which are neutralization resistant by MAb 178.1 because of an S→R or S→K mutation at position 306 (14). In addition, a group-specific neutralizing human serum neutralizes the HX10 strain very efficiently but does not neutralize the 911.32 clone (Table 1). To assess more accurately the nature of the neutralization resistance of uncloned 911.32, we molecularly cloned, after biological cloning, the 911.32 virus.

Neutralization profiles of molecular HXB2 clones chimeric for 911.32 *env* genes. A schematic representation of the envelope fragments, all cloned in a full-length HXB2 molecular clone, is shown in Fig. 1. In this way, we obtained three infectious molecular clones with different envelope regions and four infectious chimeric molecular clones with combinations of the different envelope regions. Upon transfection of these molecular clones in SupT1 or C8166 cells, the replication kinetics and the formation of syncytia for all virus clones were comparable with those for the wild-type HXB2 (data not shown). To examine neutralizing sensitivity, we tested the viral stocks of the three molecular clones (4.4, 5.1, and 6.8) against well-defined polyclonal and monoclonal antibodies (Fig. 1). Viral stock titers ranged from $10^{-3.1}$ to $10^{-2.8}$ per 40 μl. The first set of sera consisted of two autologous plasmas of chimpanzee 911. The chimpanzee plasma obtained 36 weeks after infection did not neutralize virus 5.1 but did neutralize virus 4.4 at a titer of 20 and virus 6.8 at a titer of 160. The chimpanzee plasma obtained 87 weeks after virus infection neutralized viruses 4.4 and 5.1 at reciprocal dilutions of 10 as well as virus 6.8 at a titer of 80.

The chimpanzee bulk virus isolate (911.32) is resistant to neutralization by a broadly human neutralizing serum, 1171 (Table 1). This serum neutralized virus 6.8 at a titer of 320, but again viruses 4.4 and 5.1 were neutralization resistant. To further characterize the neutralization phenotypes, we used V3- and CD4-binding-site-specific MAbs and sCD4 for confirmation. A broadly neutralizing human MAb (GP13), directed to a conserved conformation-dependent neutralization epitope that overlaps the CD4-binding site of gp120 and with neutralizing activity to both HIV-1 IIIB and MN but not RF (not shown), was used. Virus 5.1 again showed reduced neutralization sensitivity. Virus 4.4 was resistant at a titer of <5, whereas virus 6.8 was neutralization-sensitive at a titer of 400. Titers obtained with sCD4 confirmed the titers of GP13, whereas viruses 4.4 and 5.1 were less sensitive and virus 6.8 had a titer of 100. The most clear-cut difference in titers were observed with the V3-specific MAbs. Viruses 4.4 and 5.1 were neutralized only at low titers by these MAbs and were as neutralization resistant as the biological clones and the original bulk chimpanzee isolate (911) (<50; Table 1). Virus 6.8 was neutralized at a titer of 1,600 by MAb 178.1 and at a titer of 800 by MAb 5023 (different batches of MAb 5023 were used, 80 titers in Table 1 and Fig. 1 cannot be compared). These results suggest that the changes in the 2.8-kb envelope fragment of viruses 4.4 and 5.1 simultaneously affect the sensitivity for V3 antibodies as well as CD4-binding-site antibodies. Moreover, the data indicate that the 2.8-kb envelope fragment confers a neutralization phenotype to viruses 4.4 and 5.1 that is characterized by resistance to the defined set of neutralizing antibody populations present in the sera of naturally infected individuals and of experimentally infected animals.

To further characterize the neutralization phenotypes of these molecular viruses, we used (i) MAbs specific for the V3 region of IIIB and for the CD4-binding site and (ii) sCD4. We conclude that viruses 4.4 and 5.1 are neutralization resistant to V3 antibodies and less sensitive to gp120-CD4 blocking antibodies as well as to sCD4 and that virus 6.8 is neutralization sensitive to both gp120 antibodies and sCD4, in contrast to the bulk and biological viruses of chimpanzee 911. The neutralization sensitivity of virus 6.8 to monoclonal and polyclonal sera as well as to sCD4 (data not shown) is comparable to that of the HX10 isolate of HIV-1 IIIB (Table 1).

Sequence analysis of envelope determinants of neutralizing sensitivity. A 2,250-bp fragment (envelope start [nt 5771] to *Bam*HI [nt 8021]) of each of the molecular clones 4.4, 5.1, and 6.8, was sequenced. Alignments of the deduced amino acid sequences of the gene fragments relative to that of virus 5.1 are shown in Fig. 2. Sequence comparison of the three cloned gp120-gp41 envelope fragments shows that the gp120 antigenic sites recognized by sCD4 (22, 27) or by human MAb GP13 (influenced by amino acids at positions 88, 256, 257, 368, 370, 427, and 457 [26a]) as well as the linear epitopes in the V3 coding region or in sites that are reported to affect neutralization by polyclonal sera (RILAVERY; aa 579 to 586) (24, 35) were completely conserved among the clones. However, the neutralization-resistant variant 5.1 has one substitution in the V2 region, S-190→R, and two amino acid changes in gp41 in the C7 region, S-668→N and M-675→I, relative to the sensitive clone 6.8. The other neutralization-resistant variant, 4.4, has six amino acid substitutions in gp41, V-704→I, I-723→T, G-737→D, and N-750→D, as well as two gp41 amino acid changes the same as those in clone 5.1, S-668→N and M-675→I, relative to clone 6.8. Amino acid changes distin-

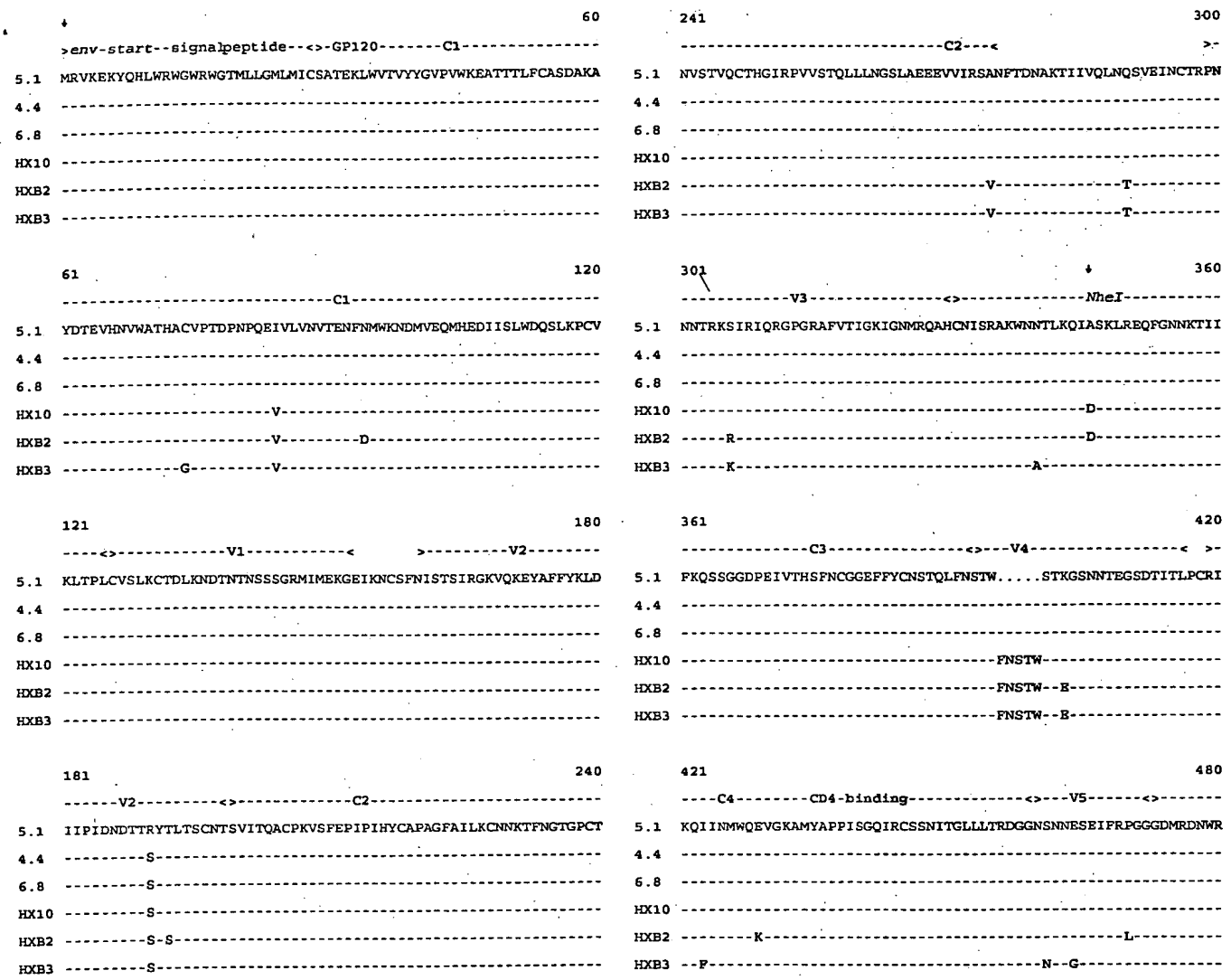


FIG. 2. Alignment of the envelope proteins and a major part of the gp41 coding regions of molecular viruses 4.4, 5.1, and 6.8 and three HIV IIIB variants, HX10, HXB2R, and HXB3, from the Los Alamos data base (17). Sequence analysis was performed by the automated sequence method as described in the text. Alignments were adjusted by eye. The predicted amino acid sequences are presented in the one-letter code. Amino acid numbers are indicated. Dashes indicate amino acid identities; points indicate gaps. The signal peptide, hypervariable region (V), constant region (C), gp120-gp41 proteolytic cleavage, *env* start, *Nhe*I, and *Bam*HI sites are indicated.

guishing these neutralization-sensitive and -resistant clones were restricted to the gp120 V2 region and the gp41 C7 region. Both the V2 region of gp120 and the second immunogenic region of gp41 (ELDKWAS; aa 662 to 668) (26, 30) are described as containing neutralization epitopes (5, 8, 12).

These results suggest that the V3 region, the CD4-binding site of gp120, and the second immunogenic region of gp41 may be domains that define the same process of entry and penetration by different means (20, 34). This process may be blocked by antibodies to continuous epitopes in each of these interactive domains or to discontinuous epitopes involving several (or all) of these interactive domains. Hwang et al. (11) and O'Brien et al. (21) have already concluded that changes in the V3 region have a direct impact on the sensitivity of a virus to sCD4 neutralization.

Exchange of gp41 and gp120 coding regions of infectious molecular clones. To further delineate the region involved in the phenotypic difference between clones 4.4, 5.1, and 6.8, exchange experiments were performed (Fig. 1). To distinguish these domains and assess their relative contribution to the neutralization phenotype as well as the possible role of the 550-bp fragment not sequenced before the envelope start site, we substituted the 3′ end (*Nhe*I-*Bam*HI; aa 346 to 752) of the *env* genes of the neutralization-resistant clones 4.4 and 5.1 for the *env* genes of the neutralization-sensitive clone 6.8 and vice versa (Fig. 1). The chimeric virus clones with the 3′ end of clone 6.8 are indeed neutralization sensitive to V3- and CD4-binding-site antibodies. The chimeric virus clones with the 3′ end of clones 4.4 and 5.1 were less neutralization sensitive than the chimeric molecular clones with the 3′ end of clone 6.8 but not as resistant as clones 4.4 and 5.1, suggesting additional contributions such as envelope structure to the

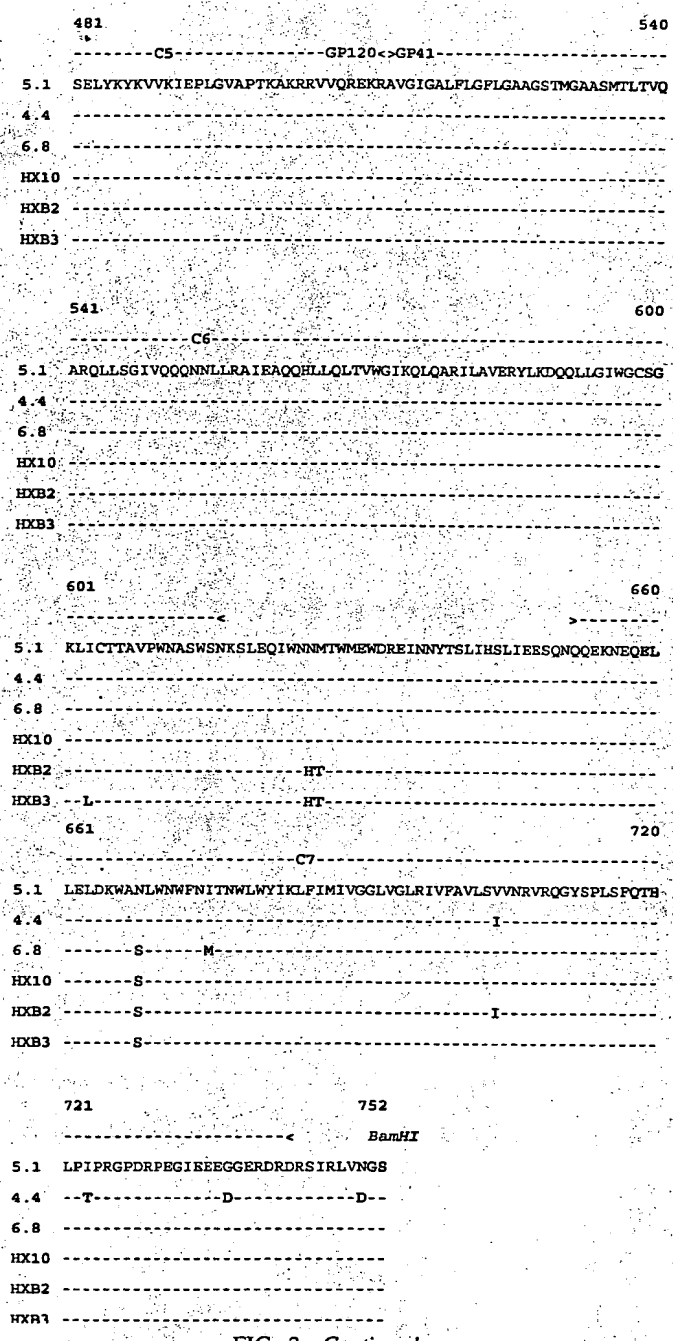


FIG. 2—*Continued.*

neutralization phenotype. Two gp41 mutations (aa 668 and 675) clearly affect the neutralization sensitivity of the viruses to V3 antibodies as well as, although to a lesser extent, the neutralizing sensitivity to gp120-sCD4 blocking antibodies and sCD4. This data set suggests again that the interaction between the gp41 region including positions 662 to 668 and V3 is stronger than with the CD4-binding site. The proposed functional interaction between the second immunogenic epitope (ELDKWAS; aa 662 to 668) of gp41 and the V3 region has not been previously described, but previous neutralization escape mutant studies have suggested that this region is not completely buried within the membrane (5).

Differences in binding of V3-specific antibodies for virion-bound gp120 may explain differences in the neutralization sensitivity of viruses. Under our assay conditions using soluble gp120, binding differences for V3 antibodies between neutralization-sensitive and neutralization-resistant viruses were not observed (data not shown).

The distinct neutralization sensitivity and the conservation of the epitopes involved suggest either that all neutralization epitopes of HIV-1 are somehow related to each other or, alternatively, that the envelope structures of these viruses are different. Third, it may be that the mutations in these clones affect the linkage between the gp120 and gp41 molecules.

The impact of the amino acid substitutions that we observed at positions 668 and 675 of the extracellular portion of gp41 on the CD4-binding site is either absent or less radical than the effect on V3 neutralization. Our results indicate that the sensitivity to neutralization by antibodies binding to gp120 can be affected not only by mutations in the gp120 coding region but also by mutations in the transmembrane protein gp41 that is associated through noncovalent interactions with gp120. These results confirm and extend previous studies by Reitz et al. (24, 35).

We thank Chitra Ramautarsing, Jacques de Jong, Wilco Keulen, Bep Klaver, and Elly Baan for technical support; Martijn Groenink and Ron Fouchier of the CLB for the envelope primers; H. Holmes of the MRC AIDS Reagent Project and contributors, Potters Bar, United Kingdom; K. Straiter of the NIAID AIDS Research and Reference Reagent Program and contributors, Bethesda, Md.; J. Sodroski and J. P. Moore for critical reading of the draft of the manuscript; and H. Rübsamen and M. Geissendorfer from the Central European Sequence Facility (Frankfurt, Germany) for help in sequencing the envelopes.

This research was sponsored in part by grant 900-502-103 from the Netherlands Organization for the Advancement of Scientific Research.

REFERENCES

1. **Albert, J., B. Abrahamsson, K. Nagy, E. Aurelius, H. Gaines, G. Nyström, and E.-M. Fenyö.** 1990. Rapid development of isolate-specific neutralizing antibodies after primary HIV-1 infection and consequent emergence of virus variants which resist neutralization by autologous sera. AIDS 4:107–112.
2. **Back, N. K. T., C. Thiriart, A. Delers, C. Ramautarsing, C. Bruck, and J. Goudsmit.** 1990. Association of antibodies blocking HIV-1 gp160-sCD4 attachment with virus neutralizing activity in human sera. J. Med. Virol. 31:200–208.
3. **Blattner, W., P. Nara, G. Shaw, B. Hahn, L. Kong, T. Matthews, D. Bolognesi, D. Waters, and R. Gallo.** 1989. Prospective clinical immunological and virologic follow up of an infected lab worker. Abstr. V Inter. Conf. AIDS, p. 510.
4. **Boom, R., C. J. A. Sol, M. M. M. Salimans, C. L. Jansen, P. M. E. Wertheim-van Dillen, and J. van der Noordaa.** 1990. Rapid and simple method for purification of nucleic acids. J. Clin. Microbiol. 28:495–503.
5. **Cao, J., L. Bergeron, E. Helseth, M. Thali, H. Repke, and J. Sodroski.** 1993. Effects of amino acid changes in the extracellular domain of the human immunodeficiency virus type 1 gp41 envelope glycoprotein. J. Virol. 67:2747–2755.
6. **de Jong, J. J., J. Goudsmit, W. Keulen, B. Klaver, W. Krone, M. Tersmette, and A. de Ronde.** 1992. Human immunodeficiency virus type 1 clones chimeric for the envelope V3 domain differ in syncytium formation and replication capacity. J. Virol. 66:757–765.
7. **Durda, P. J., L. Bacheler, P. Clapham, A. M. Jenoski, B. Leece, T. J. Matthews, A. McKnight, R. Pomerantz, M. Rayner, and K. J. Weinhold.** 1990. HIV-1 neutralizing monoclonal antibodies in-

duced by a synthetic peptide. AIDS Res. Hum. Retroviruses 6:1115–1123.
8. **Fung, M. S. C., C. R. Y. Sun, W. L. Gordon, R. S. Liou, T. W. Chang, W. N. C. Sun, E. S. Daar, and D. D. Ho.** 1992. Identification and characterization of a neutralization site within the second variable region of human immunodeficiency virus type 1 gp120. J. Virol. 66:848–856.
9. **Goudsmit, J., C. Thiriart, L. Smit, C. Bruck, and C. J. Gibbs, Jr.** 1988. Temporal development of cross-neutralization between HTLV-III B and HTLV-III RF in experimentally infected chimpanzees. Vaccine 6:229–232.
10. **Haigwood, N. L., C. B. Barker, K. W. Higgins, P. V. Skiles, G. K. Moore, K. A. Mann, D. R. Lee, J. W. Eichberg, and K. S. Steimer.** 1990. Evidence for neutralizing antibodies directed against conformational epitopes of HIV-1 gp120, p. 313–319. *In* F. Brown, R. M. Chanock, H. S. Ginsberg, and R. A. Lerner (ed.), Vaccines 90. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
11. **Hwang, S. S., T. J. Boyle, H. K. Lyerly, and B. R. Cullen.** 1992. Identification of envelope V3 loop as the major determinant of CD4 neutralization sensitivity of HIV-1. Science 257:535–537.
12. **Katinger, H., T. Muster, A. Buchacher, A. Trkola, M. Purtscher, G. Gruber, F. Steindl, G. Himmler, A. Jungbauer, and F. Ruker.** 1992. Results of preclinical screening and epitope mapping of human monoclonal antibodies against HIV-1. Abstr. 7th Cent Gardes Meet., p. 122.
13. **Langedijk, J. P. M., N. K. T. Back, P. J. Durda, J. Goudsmit, and R. H. Meloen.** 1991. Neutralizing activity of anti-peptide antibodies against the principal neutralization domain of human immunodeficiency virus type 1. J. Gen. Virol. 72:2519–2526.
14. **Langedijk, J. P. M., N. K. T. Back, E. Kinney-Thomas, C. Bruck, M. Francotte, J. Goudsmit, and R. Meloen.** 1992. Comparison and fine mapping of both high and low neutralizing monoclonal antibodies against the principal neutralization domain of HIV-1. Arch. Virol. 126:129–146.
15. **McKeating, J. A., A. McKnight, K. McIntosh, P. R. Clapham, C. A. Mulder, and R. A. Weiss.** 1989. Evaluation of human and simian immunodeficiency virus plaque and neutralization assays. J. Gen. Virol. 70:3327–3333.
16. **Montefiori, D. C., J. Zhou, B. Barnes, D. Lake, E. M. Hersh, Y. Masuho, and L. B. Lefkowitz, Jr.** 1991. Homotypic antibody responses to fresh clinical isolates of human immunodeficiency virus. Virology 182:635–643.
17. **Myers, G., B. Korber, J. A. Berzofsky, R. F. Smith, and G. N. Pavlakis (ed.).** 1992. Human retroviruses and AIDS 1992. Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, N.Mex.
18. **Nara, P. L., W. G. Robey, L. O. Arthur, D. M. Asher, A. V. Wolff, C. J. Gibbs, Jr., D. C. Gajdusek, and P. J. Fischinger.** 1987. Persistent infection of chimpanzees with human immunodeficiency virus: serological responses and properties of reisolated viruses. J. Virol. 61:3173–3180.
19. **Nara, P. L., L. Smit, N. Dunlop, W. Hatch, M. Merges, D. Waters, J. Kelliher, R. C. Gallo, P. J. Fischinger, and J. Goudsmit.** 1990. Emergence of viruses resistant to neutralization by V3-specific antibodies in experimental human immunodeficiency virus type 1 IIIB infection of chimpanzees. J. Virol. 64:3779–3791.
20. **Neurath, A. R., N. Strick, and S. Jiang.** 1992. Synthetic peptides and anti-peptide antibodies as probes to study interdomain interactions involved in virus assembly: the envelope of the human immunodeficiency virus (HIV-1). Virology 188:1–13.
21. **O'Brien, W. A., Y. Koyanagi, A. Namazie, J.-Q. Zhao, A. Diagne, K. Idler, J. A. Zack, and I. S. Y. Chen.** 1990. HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 outside the CD4-binding domain. Nature (London) 348:69–73.
22. **Olshevsky, U., E. Helseth, C. Furman, J. Li, W. Haseltine, and J. Sodroski.** 1990. Identification of individual human immunodeficiency virus type 1 gp120 amino acids important for CD4 receptor binding. J. Virol. 64:5701–5707.
23. **Reiss, P., J. M. A. Lange, A. de Ronde, F. de Wolf, J. T. Dekker, C. Debouck, and J. Goudsmit.** 1990. Speed of progression to AIDS and degree of antibody response to accessory gene products of HIV-1. J. Med. Virol. 30:163–168.
24. **Reitz, M. S., Jr., C. Wilson, C. Naugle, R. C. Gallo, and M. Robert-Guroff.** 1988. Generation of a neutralization-resistant variant of HIV-1 is due to selection for a point mutation in the envelope gene. Cell 54:57–63.
25. **Robert-Guroff, M., M. Brown, and R. C. Gallo.** 1985. HTLV-III-neutralizing antibodies in patients with AIDS and AIDS-related complex. Nature (London) 316:72–74.
26. **Robinson, W. E., Jr., M. K. Gorny, J.-Y. Xu, W. M. Mitchell, and S. Zolla-Pazner.** 1991. Two immunodominant domains of gp41 bind antibodies which enhance human immunodeficiency virus type 1 infection in vitro. J. Virol. 65:4169–4176.
26a. **Schutten, M., A. McKnight, R. Huisman, A. C. Andeweg, M. Thali, J. A. McKeating, J. Sodroski, J. Goudsmit, and A. D. M. E. Osterhaus.** 1993. Further characterization of an antigenic site of HIV-1 gp120 recognized by virus neutralizing human monoclonal antibodies. AIDS 7:919–923.
26b. **Sodroski, J.** Personal communication.
27. **Thali, M., C. Furman, D. D. Ho, J. Robinson, S. Tilley, A. Pinter, and J. Sodroski.** 1992. Discontinuous, conserved neutralization epitopes overlapping the CD4-binding region of human immunodeficiency virus type 1 gp120 envelope glycoprotein. J. Virol. 66:5635–5641.
28. **Thiriart, C., M. Francotte, C. Cohen, C. Collignon, A. Delers, S. Kummert, C. Molitor, D. Gilles, P. Roelants, F. van Wijnendaele, M. de Wilde, and C. Bruck.** 1989. Several antigenic determinants on the gp120 moiety of HIV-1 gp160 are hidden on the mature gp120. J. Immunol. 143:1832–1836.
29. **Tremblay, M., and M. A. Wainberg.** 1990. Neutralization of multiple HIV-1 isolates from a single subject by autologous sequential sera. J. Infect. Dis. 162:735–737.
30. **Tyler, D. S., S. D. Stanley, S. Zolla-Pazner, M. K. Gorny, P. P. Shadduck, A. J. Langlois, T. J. Matthews, D. P. Bolognesi, T. J. Palker, and K. J. Weinhold.** 1990. Identification of sites within gp41 that serve as targets for antibody-dependent cellular cytotoxicity by using human monoclonal antibodies. J. Immunol. 145: 3276–3282.
31. **Von Gegerfelt, A., J. Albert, L. Morfeldt-Månson, K. Broliden, and E.-M. Fenyö.** 1991. Isolate-specific neutralizing antibodies in patients with progressive HIV-1-related disease. Virology 185:162–168.
32. **Wahlberg, J., J. Albert, J. Lundeberg, A. Von Gegerfelt, K. Broliden, G. Utter, E.-M. Fenyö, and M. Uhlén.** 1991. Analysis of the V3 loop in neutralization-resistant human immunodeficiency virus type 1 variants by direct solid-phase DNA sequencing. AIDS Res. Hum. Retroviruses 7:983–990.
33. **Weiss, R. A., P. R. Clapham, R. Cheingsong-Popov, A. G. Dalgleish, C. A. Carne, I. V. D. Weller, and R. S. Tedder.** 1985. Neutralization of human T-lymphotropic virus type III by sera of AIDS and AIDS-risk patients. Nature (London) 316:69–72.
34. **Willey, R. L., E. K. Ross, A. J. Buckler-White, T. S. Theodore, and M. A. Martin.** 1989. Functional interaction of constant and variable domains of human immunodeficiency virus type 1 gp120. J. Virol. 63:3595–3600.
35. **Wilson, C., M. S. Reitz, Jr., K. Aldrich, P. J. Klasse, J. Blomberg, R. C. Gallo, and M. Robert-Guroff.** 1990. The site of an immune-selected point mutation in the transmembrane protein of human immunodeficiency virus type 1 does not constitute the neutralization epitope. J. Virol. 64:3240–3248.